

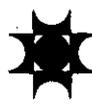
No 0000198



1988

84887

**Report on the
Battle Creek Health Study**

 **Michigan Department of Public Health**

**Report on the
Battle Creek Health Study**

by

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

and

**Arthur W. Bloomer, MS
Center for Environmental Health Sciences
Michigan Department of Public Health
Lansing, Michigan**

1988

This study was largely supported by funds from the Comprehensive Environmental Response, Compensation, and Liability Act trust fund, by Interagency Agreement with the Agency for Toxic Substances and Disease Registry, U.S. Public Health Service, Department of Health and Human Services.

Report Editing and Production

**R. E. Mosher, PhD
G. E. Beane, PhD
MB Research Associates**

Acknowledgements

This report is the result of considerable effort by many individuals and organizations plus the expenditure of an estimated \$1,250,000 in Michigan and Federal funds. In this regard, I wish to acknowledge and express my thanks and gratitude to the many agencies and persons identified below. Without their sincere interest, concern and commitment this environmental epidemiological study of human exposure to volatile organics in drinking water would not have been possible.

I want to officially thank the Centers for Disease Control, Center for Environmental Health and Injury Control for their willingness to work cooperatively with this Department in conducting this study.

The financial support of the Agency for Toxic Substances and Disease Registry was vital in funding the major portion of this study.

The assistance from Michigan's Governor James J. Blanchard and Congressman Howard Wolpe were essential in obtaining the monies and other support necessary.

The cooperation of the Calhoun County Health Department's staff, the Calhoun County Board of Health, the Cass County Health Department's staff and the Cass County Board of Health helped in identifying the geographical areas of concern and in the collection of necessary water specimens for analysis.

The staff of Western Michigan University's Kercher Center for Social Research conducted the field and data gathering phases of this study in an exemplary manner. Their subcontractor, the Borgess Medical Center conducted their portion of the study in a dependable and professional way.

The staff of the Michigan Department of Public Health's Center for Environmental Health Sciences, Bureau of Laboratory and Epidemiological Services and the Center for Health Statistics all performed vital roles in the conduct of this study.

The cooperation of the medical care communities (personal physicians and hospitals) was outstanding. Their sharing of the medical records under their control was a vital component of verifying diagnoses and illnesses.

Of course, this study could not have been conducted without the excellent cooperation of the many citizens who were the subject of concern. It was only through each individual's willingness to be part of the investigation that we could even begin to consider this undertaking. They made time available in their busy lives to allow this work to happen. I personally want to thank them for their efforts and cooperation.

Cordially,



Raj M Wiener
Acting State Health Director

Chapter	Contents	Page
0	EXECUTIVE SUMMARY	1
1	INTRODUCTION	13
	1.1 The initial situation	13
	1.2 The contaminants	14
	1.3 The sites	15
	1.4 Phase I	18
	1.5 Phase II	18
	1.6 Persons and institutions involved	19
	1.7 Contents of the report	20
	1.8 Peer review	21
2	SUMMARY OF THE METHODOLOGY	23
	2.1 Introduction	23
	2.2 Statistical power as a basis for study size	23
	2.3 Phase I	24
	2.4 Phase II	27
	2.5 Statistical analysis	34
3	TOXICOLOGY	35
	3.1 Introduction	35
	3.2 Target tissues and systems	35
	3.3 Mutagenicity and cancer	36
	3.4 Toxicity in specific exposure scenarios	37
	3.5 LD ₅₀ , LOEL, and NOEL	39
4	EXPOSURE ASSESSMENT	43
	4.1 Introduction	43
	4.2 Estimating individual exposure to VOCs	43
	4.3 Conversion of exposure to dose	48
	4.4 VOCs in drinking water of comparison areas	48
	4.5 VOCs in water supplies of former residents	48
	4.6 VOC-exposure through bathing and showering	49
	4.7 VOC-exposure from secondary sources	50
	4.8 Exposure to other chemicals	50
5	HEALTH ASSESSMENT	53
	5.1 Introduction	53
	5.2 The questionnaire	53
	5.3 Medical records	54
	5.4 Clinical examination	54
	5.5 Validation of health data	55
	5.6 State-generated mortality and hospital discharge rates	56
6	EPIDEMIOLOGIC ANALYSIS	57
	6.1 Introduction	57
	6.2 Followup time	57
	6.3 Lag time	57
	6.4 Validation of health data	58
	6.5 Eligibility of cases for analysis	58
	6.6 Univariable analysis of exposure-effect associations	58
	6.7 Multivariable analysis	59
	6.8 Overview of statistical analyses	61
	6.9 Incidence rate and prevalence	62
	6.10 Evaluation of the results of statistical analyses	63

Chapter	Page
7 QUALITY CONTROL FOR CLINICAL LABORATORIES	65
7.1 Overview	65
7.2 Serum chemistry profile	65
7.3 Hematology profile	66
7.4 Urinary profile	66
7.5 Serum PCB and PBB	67
8 RESULTS	69
8.1 Introduction	69
8.2 PCB, PBB, and chlorinated pesticides	69
8.3 Clinical examination	71
8.4 Diabetes mellitus	73
8.5 Hypertension	74
8.6 Gallbladder disease	74
8.7 Abnormal pregnancy outcomes	75
8.8 Congenital defects	76
8.9 Thyroid disease	78
8.10 Epilepsy	79
8.11 Ulcer	81
8.12 Kidney disease	82
8.13 Cancer	83
8.14 Skin diseases and allergies	88
8.15 Other diseases and conditions	90
8.16 Effect of tracking on disease frequency	91
9 DISCUSSION	95
9.1 Introduction	95
9.2 Exposure assessment	95
9.3 Health assessment	100
9.4 Statistical analysis	101
9.5 Results	104
9.6 Study findings unrelated to VOC Exposure	108
9.7 Summary	111
REFERENCES	113
APPENDIX A Questionnaire	123
APPENDIX B Contractor's report	145
APPENDIX C Groundwater modeling	157
APPENDIX D State-generated mortality and hospital discharge rates	179
APPENDIX E Diabetes mellitus	199
APPENDIX F Hypertension	215
APPENDIX G Gallbladder disease	231
APPENDIX H Abnormal pregnancy outcomes	245
APPENDIX I Algorithm for case eligibility and calculating TAE	257

LIST OF ABBREVIATIONS USED IN THE REPORT AND ITS APPENDICES

Abbreviation	Meaning	Reference Chapter or Appendix
ADI	allowable daily intake	3
AADI	adjusted allowable Daily intake	3
b or b ₁	slope or regression coefficient in C = f(T)	4
BUN	blood urea nitrogen level	8
C	concentration of a chemical in water (in ppb)	4
C _s	chemical concentration in a particular sample	4
C = f(T)	chemical concentration as a function of time	4, App. C
CDC	Centers for Disease Control, Atlanta, GA	
CEH	Center for Environmental Health, organizational unit within the CDC, now renamed CEHIC	
CEHIC	Center for Environmental Health and Injury Control, new name of CEH, part of the CDC	
CIS	cis-1,2-dichloroethylene, one of the VOCs	2, 3
CT-curve	concentration - time curve	4, App. C
CT-model	linear model predicting concentration from time	4, App. C
11DCA	1,1-dichloroethane, one of the VOCs	2, 3
12DCA	1,2-dichloroethane or ethylene dichloride, a VOC	2, 3
DCE	1,1-dichloroethylene, one of the VOCs	2, 3
DDT	dichloro-diphenyl-trichloroethane, a pesticide	8
DOSE	expression for the orally absorbed amount of VOCs	4, 6
DOSEVOC	same as DOSE, specially referring to VOCs	
DOSEVCL	same as DOSE, for absorbed VOCs + chloroform	
g	gram, 1/1000 of a kilogram	
GGTP	serum level of gamma glutamyl transpeptidase	8
HbA1C	(glycosylated) hemoglobin A ₁ C	App. E
IR	incidence rate, in particular population-based	
kg	kilogram	
LDH	serum level of lactate dehydrogenase	8
LOEL	lowest observed effect level, also covering adverse effect	3
Log	natural logarithm	
LogC = T model	linear model predicting log concentration from time	4, App. C
m	meter	
MDPH	Michigan Department of Public Health	
NOEL	no observed effect level (also covering adverse effects)	3, 9
OGTT	oral glucose tolerance test	App. E
OR	odds ratio	6
PBB	polybrominated biphenyls	8
PCB	polychlorinated biphenyls	8
PCE	perchloroethylene or 1,1,2,2-tetrachloroethylene one of the VOCs	2
ppb	parts per billion (concentration measure)	
ppm	parts per million (concentration measure)	
QA/QC	quality assurance and quality control	
QI	Quetelet index, equal to weight(kg) / height(m) ²	8

Abbreviation	Meaning	Reference Chapter or Appendix
RR	relative risk or (incidence) rate ratio	6
SGOT	serum glutamic oxalo-acetic transminase	8
SGPT	serum glutamic pyruvic transaminase	8
T	time in number of months since January 1, 1970. Is used in T ₁ , T _s , TDEATH, TDIAG, TFU1, TFU2, TIN, TOUT, TQUEST, and TSTOP	4, App. C
T _s	date a particular water sample was taken	4, App. C
T ₁	date that a well became contaminated at the detection level of 1 ppb	4, App. C
TAE	total accumulated exposure, general expression	4, 6
TAEVOC	same as TAE, in particular for VOCs	4, 6
TAEVCL	same as TAE, in particular for VOCs + chloroform	4, 6
TCA	1,1,1-trichloroethane, one of the VOCs	2, 3
TCE	trichloroethylene, one of the VOCs	2, 3
TDEATH	date of death	App. I
TDIAG	first date of diagnosis of a disease	6
TFU	time (duration) of followup	4, 6
TFU1	start of followup period, equal to 1-1-1970 or TIN	4, 6
TFU2	end of followup period, equal to date of interview, death, or (if a disease developed) TDIAG	4, 6
TIN	date a person moved into to the study area	4, 6
TOUT	date a person moved out of the study area	4, App. I
TQUEST	date a person was interviewed	App. I
TSTOP	date a participant from the exposed cohort stopped drinking contaminated water.	4, 6
VOC(s)	volatile organic chemical(s)	
WASH	measure of water use for bathing and showering	4, 6
WASHVOC	WASH times sum of concentrations of VOCs in the current water sample	6
WASHVCL	same as WASHVOC for sum of VOCs + chloroform levels	6
WATER	amount of unheated tapwater consumed at home	4

Executive Summary

0.1 Background

In 1981, the aquifer serving the Battle Creek area, Calhoun County, Michigan, was found to be contaminated with volatile organic chemicals (VOCs). The most seriously affected city wells were closed down, and residents in the adjacent Verona Park, whose private wells were contaminated, were offered bottled water and municipal shower facilities until their homes were connected to the city water mains. In 1986, new city wells north of the city's well field (the Verona well field) came into operation.

Groundwater contamination with the same VOCs was later detected in Springfield, southwest and adjacent to Battle Creek, and in Dowagiac, Cass County. The source of contamination was identified in Battle Creek and Dowagiac, but not in Springfield. Responding to public concerns, the Michigan Department of Public Health (MDPH) and the U.S. Centers for Disease Control (CDC) conducted a comprehensive epidemiologic study of the potential health effects of the groundwater contamination. This report describes the rationale, design, implementation, and the results of the study.

0.2 Rationale

A review of the relevant literature revealed that adverse effects of chronic exposure to VOCs had been observed only at levels much higher than had been observed in drinking water. However, this conclusion was reached on the basis of studies reporting on exposures to a single chemical. There have been no well-designed and carefully analyzed studies of multiroute exposure to multiple VOCs. To obtain a firm basis of facts for future risk management, and to properly address the public's concerns, a comprehensive epidemiologic study was carried out. This study specifically addressed the adverse effects of concomitant exposure to the seven VOCs detected in the groundwater, utilizing state-of-the-science knowledge and methods.

0.3 Study Design

The Battle Creek Health Study was designed as a retrospective cohort study. In this design, a cohort of exposed people is compared with a reference cohort of unexposed people with regard to the incidence of diseases or other health parameters during a defined period of observation (followup period). The data are then analyzed to determine whether or not exposure to VOCs could explain possible differences in disease incidence. To optimize the conditions for valid conclusions, the study design and analysis followed the principles of risk assessment for hazardous chemicals, as summarized below:

- 1) The exposure assessment must include exposure to all VOCs by oral, skin, and inhalation routes of entry, and be quantified in terms of dose and duration.
- 2) Quality control procedures must apply to both exposure and health data.
- 3) Exposure incurred after a disease is diagnosed, and a disease diagnosed prior to exposure, must be discounted prior to estimating the risk.
- 4) The risk must be evaluated in the context of a dose-response analysis.

- 5) The analysis must properly control for the effect of confounding factors (to compensate for the fact that people cannot be experimentally subjected to exposure).
- 6) The effect of the methods and of uncertainties in the data on the outcome of the analysis must be evaluated.

0.4 Formation of Study Cohorts

The exposed cohort included current and former residents of dwellings with contaminated wells in Verona Park, Springfield, and Dowagiac. The reference neighborhoods were selected for comparability to the contaminated areas with respect to the age, size, and value of the dwellings. Demographic data compiled in Phase I of the study were used for screening reference households for Phase II (the actual epidemiologic study) in terms of comparability of age and sex distribution, and the year of moving into the study area. Once a household was selected to be part of the final cohort, all former residents of the dwelling were traced to be included in the Phase II study. About 20% of the eligible exposed population, and 30% of the reference population refused to participate in Phase II. In total, a cohort of 251 potentially exposed people and a reference cohort of 498 people took part in the study. The cohorts were quite comparable with regard to age, education and income level, occupational exposure to VOCs and other chemicals, the year of moving to the study area, and the male/female ratio.

0.5 Quality Control

No quality control and assurance procedures have been defined for data other than laboratory tests. Therefore, this study utilized a multisource approach to gathering health information to optimize the quality of the health data. Data from the multiple sources were compared to validate disease diagnoses, and to improve the accuracy of the estimate of the date when a disease was first diagnosed. In the exposure assessment, testing of the drinking water supply for VOCs was extended to the current water supply of persons who had moved out of the study area during the observation period. The quality of the transfer of raw data to computer tapes was checked by the contractor and, in a later phase, by the CDC. A rigorous quality control and assurance procedure was established for laboratory tests of water, blood, and urine specimens, and the blood pressures were measured with a mechanical device which avoided reader bias.

0.6 Toxicity of the VOCs

The seven VOCs have almost identical toxicity profiles. All are known to have an effect on the central nervous system, the liver, and the kidney, and may cause heart arrhythmias. High concentrations may have a local effect on mucous membranes and skin. Some VOCs are suspected to be animal carcinogens, but no evidence of human carcinogenicity has been found. Since VOCs can pass the placenta, VOCs at sufficiently high levels may also have an effect on the fetus.

0.7 Exposure Assessment

In virtually all environmental studies, the quality of historical exposure data is extremely poor, since the data have been compiled routinely without a specific study in mind. In the Battle Creek study, the only data on historical exposure consisted of the test results of a recent (one-time) sampling of private wells in the contaminated area. A mathematical model was

developed from existing monitoring data in the Battle Creek city wells to estimate $C=f(T)$, the change in VOC concentrations with time. This $C=f(T)$ was then extrapolated to the residential wells in the Verona Park area to estimate when the contamination of a particular well started (T_1), and to calculate for individuals the total accumulated exposure (TAE) for single and multiple VOCs in any combination. Interview data complemented this information, enabling the adjustment of the TAE for the actual period of time that people lived in the study area, and the conversion of TAE into DOSE (equal to TAE times the volume of unheated tap water consumed). No monitoring data were available to estimate $C=f(T)$ in Dowagiac and Springfield. The exposure assessment for these sites was necessarily based on the assumption that VOCs were present at a constant level since 1970. A composite value TAEVOC (and DOSEVOC) was calculated, equal to the sum of the VOC-specific TAEs (based on the principle of dose-equivalents).

Chloroform is formed in city water as a byproduct of chlorination, and its toxicologic profile is identical to that of the VOCs of interest. Exposure to chloroform must, therefore, be taken into account in an alternative analysis of the possible effects of exposure to VOC-contaminated water. Exposure values TAEVCL and DOSEVCL (the VOCs combined with chloroform) were computed as alternative exposure expressions.

Table 0.1 Conventional (but erroneous) exposure values, expressed as the sum of VOC-specific concentrations in the current water sample (sum VOCs), and the sum of the exposure accumulated during the period of residence in the exposed area (TAEVOC).

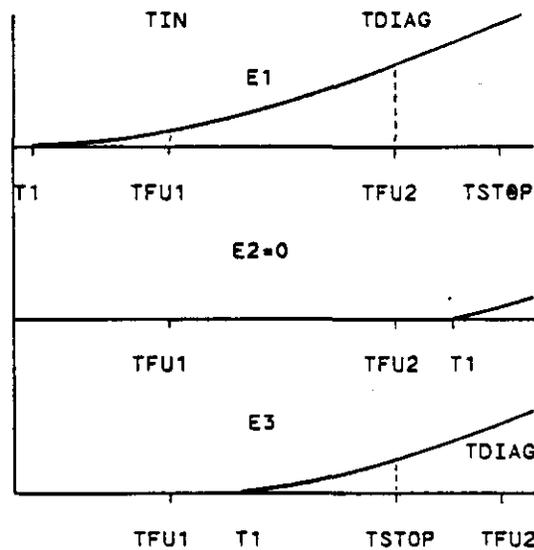
Current Water Sample			Concentration-Time Integration		
People	%	Sum VOCs (ppb)	People	%	TAEVOC (ppb-months)
0	0	0	28	11	0
138	55	1 - 9	49	20	1 - 99
60	24	100 - 1,000	33	13	100 - 1,000
43	17	1,000 - 10,000	58	23	1,000 - 10,000
10	4	10,000 - 100,000	47	19	10,000 - 100,000
-	-	-	25	10	100,000 - 1,000,000
-	-	-	11	4	1,000,000 - 10,000,000

Table 0.1 clearly illustrates the effect of a conventional, but erroneous, exposure value, equal to the simple sum of the VOC-concentrations in a current water sample. This value ignores the fact that the wells have not always been contaminated. In fact, using all available data (concentration-time integration), it was shown that 11% of the residents of the exposed area had never been exposed, because they had moved to the study area before T_1 ; TAEVOC as the exposure value resulted in a substantially different distribution of exposure levels.

Figure 0.1 shows the complex relationship between time-events, changes in exposure level over time, and the VOC-specific TAE. The horizontal axis in this figure is the time-of-followup (TFU), the vertical axis represents the concentration C of the VOC of interest. The first scenario is of a person (#1) with a disease diagnosed at date TDIAG, who moved to the area at date TIN and stopped drinking VOC-contaminated water at date TSTOP. The TFU of this person equals the period of time between TFU1 (equal to TIN) and TFU2 (equal to TDIAG), and TAE for this person and this VOC equals the shaded area under the curve. The well of person #2 became contaminated at date T_1 after TFU2 (now equal to the date of interview, since no disease had developed) rendering the exposure equal to zero. For person #3, T_1 was after TIN,

but before a disease was diagnosed. As this individual stopped consuming contaminated water at TSTOP before TDIAG, TAE was calculated from T_1 to TSTOP.

A computer algorithm was developed to facilitate the estimation of TAE for each of the many possible scenarios of time events, VOCs, and individuals. A major problem was the calculation of TAE and DOSE for the VOCs absorbed through the skin and the lungs. Experimental studies which should have provided the necessary absorption rates were not standardized. This made their results unusable for estimating the composite multiroute TAE and DOSE with a reasonable degree of accuracy. Therefore, other-than-oral VOC exposure was treated as a covariate in a multivariable model.



TIN = date subject moved to study area.

T_1 = date the well of this person became contaminated with a specific VOC.

TFU1, TFU2 = date followup began and date it stopped.

TSTOP = date use of water stopped.

E1 : shaded area from TFU1 (=TIN) to TFU2 (=TDIAG) represents the TAE for this person and this VOC.

E2 : TAE = 0 because TDIAG was later than T_1 .

E3 : the TAE is the shaded area from T_1 to TSTOP, which was before a disease was diagnosed (=TDIAG).

Figure 0.1 Determinants of scenario- and VOC-specific total accumulated exposures (TAEs) for three persons with different time events.

0.8 Health Assessment

Health information was obtained through interviews, information on the use of prescription or over-the-counter drugs, clinical examinations, and abstracts of medical records from hospitals and private physicians. In the questionnaire, multiple questions addressed the same issue, to improve the accuracy of the responses, and to identify discrepancies. The information from these sources were then pooled to generate validated disease-specific data bases.

Validation of the health information utilized all available information. There were large discrepancies between interview responses and medical records in terms of the dates of diagnosis and the types of diseases. Months to years may have elapsed from the time when the toxic effect of exposure became irreversible to the time when the disease became detectable and was eventually diagnosed. Therefore, in the analysis, a provision was made to account for the time lag between the estimated date the disease was initiated (the date essential to evaluate the risk of exposure), and the earliest date of diagnosis.

The health information revealed some potential sources of public health concern. Many people were unaware that they had a disease requiring medical treatment. Many others knew about their disease, but showed no evidence of current treatment, demonstrated poor compliance to the therapy, or might have been over-treated. It is possible that some people might have had difficulties in obtaining access to the community health system. The data also suggest that the quality of communication between physicians and patients can be improved. This study was not designed to investigate these areas of public health concern, however.

0.9 Statistical Analysis

In a conventional analysis, the effect of exposure is measured as the odds ratio (OR) (from fourfold tables of the disease status by exposure status), or as the relative risk (RR), in which the time of followup (TFU) is the denominator of the risk. A better approach is a multivariable analysis using a proportional hazard model. This model assumes a constant exposure level during the followup period. Since exposure levels did vary over time, and most exposure stopped long before the end of the followup period, modification of the model was required. For this purpose, the followup period was partitioned into periods of one year, in which each individual had a period-specific exposure and health status. A final data set was then created by pooling the period-specific observations. The analysis was done on this pooled data set, blocking for the period.

This model could not be used for the analysis of pregnancy outcomes, as the followup period does not end at the termination of the pregnancy. Therefore, a logistic regression model was used with the pregnancy as the unit of observation.

Multivariable models require a sufficient number of cases for each disease, relative to the number of variables in the model. Since the models would include a variable for oral exposure (TAE or DOSE) and a variable for use of water for bathing and showering (WASH), at least 20 cases would be required for a model that included even one confounding factor. Thus, multivariable analysis was limited to diabetes, hypertension, gall bladder disease, and some pregnancy outcomes, since these were the only conditions with at least 20 cases.

In the evaluation of the results from the analysis, the p-value (a measure of the probability of a chance effect) was given less weight than consistency of results and a positive dose-response, since p-values were likely to be insignificant due to the small number of cases. To judge the consistency of the results, the analysis was repeated for at least five exposure expressions, two populations (the total population and a population limited to the Verona exposed plus the Calhoun County reference cohorts), and by gender (for hypertension and cancer). In multivariable models, each exposure expression was calculated in three different ways. This very unusual approach was chosen to provide the means to judge for consistency and to avoid errors which might result from selecting a single analytical model.

0.10 Results

The following is a summary of the results of the analysis of possible associations of exposure to VOCs with: exposure to other exogenous chemicals; clinical tests; and diseases occurring frequently enough to warrant statistical analysis.

0.10.1 Exogenous chemicals

Dichloro-diphenyl-trichloroethane (DDT): Residents of the exposed area tested slightly *less* frequently positive for DDT or its metabolites than residents of reference neighborhoods (97.7% versus 98.7%) and had a *lower* median serum level (4.1 ppb versus 4.5 ppb).

Heptachlor and chlordane: Only 2.9% of the tests (too few for an analysis) yielded serum values above the detection limit with a maximum of 5 ppb.

Polybrominated biphenyls (PBB): Exposure to this chemical is specific to Michigan. Residents of the exposed area tested *more* frequently positive for PBB than residents of reference neighborhoods (31.2% versus 28.7%), and they had a slightly *higher* median serum level (1.8 ppb versus 1.5 ppb).

Polychlorinated biphenyls (PCB): Residents of the exposed area tested *less* frequently positive for PCB than residents of the reference areas (40.4% versus 46.0%) and had a *lower* median serum level (5.5 ppb versus 5.9 ppb).

The differences in the serum levels of these exogenous chemicals are statistically insignificant; there is no evidence that exposure to VOCs has increased the uptake of these compounds. It is unknown what the health effects may be of these levels of these compounds. An in-depth analysis of the already available data might yield useful information on this point. No such analysis was done due to time constraints, and because it was not relevant to VOCs.

0.10.2 Clinical Chemistry

Correlation tests for continuous variables showed that, with a few exceptions, all clinical tests were *negatively* correlated with exposure. In the few cases in which the coefficient was positive, the r-value was extremely small ($r < 0.1$) and the p-value was larger than 0.1. In agreement with this finding, abnormal test values (a value is abnormal if it exceeds the limit of the normal range) were much *less* frequent in the exposed cohort than in the reference cohort. The *deficit* of abnormal values among the exposed was often significant at $p < 0.1$.

0.10.3 Diseases

Table 0.2 gives a summary of the frequencies of the diseases which have been analyzed. The number of cases refer to all cases encountered, regardless of the date of diagnosis, and whether the diseases were detected while the individuals were living in the study area or not. Thus, these numbers are what would have been the result if this study were a simple frequency survey, and they represent the popular perception of a disease frequency. Differences in frequencies between the cohorts always occur. A low (lower than 0.1) p-value for the *ratio* of the frequency rates indicates that these differences are statistically significant. The table shows that disorders with a lower or a higher rate of occurrence in the contaminated area are equally distributed; this is the result expected if VOCs had no adverse health effect at the prevailing exposure level.

If the comparison is limited to the six statistically significant differences, five out of six disorders have a higher rate for residents of the contaminated area. This is more than might be expected if chance were the only factor determining differences in rates. This may raise the question of whether these differences could have been caused by exposure to VOCs. To answer this question, however, the comparison of crude numbers is not the proper approach, since many cases were diagnosed before exposure began, and the individual exposure status was not assessed.

Table 0.2 Overview of disease frequencies, the number of cases is *unrestricted*, as no eligibility criteria were applied. Therefore, the rate/1000 population is *not* an incidence or a prevalence rate! The p-values were derived from a one-sided Fisher's exact test for the *ratio* of the rates.

Disease		Contaminated area		Reference area		p-value	
		cases	rate/1000	cases	rate/1000		
Diabetes	*	13	51.8	15	30.1	0.10	
Hypertension	*	33	131.5	82	164.7		
Gallbladder	*	15	59.8	19	38.2		
Hypothyroidism		6	23.9	14	28.1		
Epilepsy		4	15.9	10	20.1		
Kidney disease		6	23.9	12	24.1		
Stomach ulcer	*	7	27.9	15	30.1		
All cancers in men	*	5	42.0	10	43.5		
All cancers in women	*	11	83.3	11	41.0		0.07
Birth defects	**	8	33.5	19	39.5		
Miscarriage	**	46	161.4	42	80.3	<0.001	
Prematurity	**	8	33.5	21	43.7		
Stillbirth	**	5	20.9	8	16.6	0.001	
Low birth weight	**	18	76.3	41	85.8		
Skin - hives		23	91.6	32	64.3	0.02	
- rashes		21	83.7	14	28.1		
- acne		5	19.9	27	54.2		
- psoriasis		11	43.8	11	22.1		0.08
- eczema		31	123.5	42	84.3		
Specific skin allergies		21	83.7	48	96.4		
Asthma or hay fever		31	123.5	53	106.4		

* The available information was sufficient for case validation, and the case-definition followed the diagnostic categories (second column). The categories are explained in the Appendices and, for cancer and peptic ulcer, in Chapter 8. For other disorders, the case validation was limited or (as for skin diseases and allergies) not possible at all.

** The denominator for calculation of the miscarriage rate is the number of all pregnancies; the denominator for the calculation of the birth defect, prematurity, stillbirth, and low birth weight rates is the number of pregnancies *not* ending in miscarriage. For birth weight, the denominator is somewhat smaller due to missing values.

The required corrections are: 1) *delete cases diagnosed before the individual moved to the study area*, (these cases could not possibly be attributed to the VOCs); and 2) *accept cases only if diagnosed in 1970 or later*. The reason for this criterion of eligibility is that the search for medical records was limited to the period 1970-1985. A search for records covering the entire lifespan (which may go back as far as 1895) was obviously not feasible. Thus, information about diseases diagnosed prior to 1970 was based entirely on interview responses. This source of data has a high likelihood of being inaccurate or biased towards better disease recall by people in the exposed area. If these two criteria are applied, the pattern of differences in disease rates changes, as shown in Table 0.3.

In this table, the relation between disease occurrence and exposure to VOCs is depicted in the odds ratios (OR) and the relative risk (RR). The diseases are divided into three groups. The first of these groups consists of disorders with a sufficient quality of data and number of cases for multivariable analysis. The second group comprises disorders of which the number of cases was too small, and the quality of health data too poor to warrant multivariable analysis. The third group includes skin disorders and allergies, disorders *not* validated for the correctness or date of the diagnosis because of the lack of a supporting medical record in the majority of cases.

In Table 0.3, exposure is expressed in four alternative ways, each with its own rationale. DOSE is the better measure of exposure, since TAE merely refers to the availability of VOCs and not the consumed amount. DOSEVCL takes into account that the exposure to VOCs should include chloroform (CHL) because this compound has the same toxicology profile as the other VOCs. DOSEVCL is, thus, conceptually the best exposure expression, but it requires information on the volume of unheated tap water consumed at home. This information may be unreliable, and it was missing for some people, which reduced the number of cases in the final analysis. TAE may, therefore, be a suitable alternative exposure expression if one does not want to accept the uncertain quality of the data or the smaller number of cases involved in using DOSE. However, the use of TAE assumes that people have the same pattern of water intake across cohorts, an assumption shown to be incorrect.

Table 0.3 Overview of odds ratios (OR) and rate ratios (RR) for dichotomous exposure expressions. An asterisk indicates statistical significance at $p < 0.1$ (one-sided Fisher's exact test). For details, see Chapter 8 and the disease-specific appendices.

Disease and number of cases eligible for analysis ()	Exposure = VOCs only		Exposure = VOCs + CHL		TAEVCL		DOSEVCL	
	TAEVOC OR	DOSEVOC RR	TAEVOC OR	DOSEVOC RR	OR	RR	OR	RR
Diabetes (20)	1.99*	2.06*	2.60*	2.77*	0.73	0.86	1.04	1.21
Hypertension (92)	0.49*	0.54*	0.49*	0.56*	0.55*	0.66*	0.59*	0.70*
Gallbladder disease (28)	0.85	0.89	0.91	0.97	0.34*	0.41*	0.41*	0.48*
Cancer men (13)	1.04	1.12	1.36	1.51	0.42	0.53	0.41	0.53
Cancer women (17)	1.73	1.71	1.64	1.67	1.55	1.66	1.48	1.58
Epilepsy (6)	0.47	0.49	0.56	0.61	2.44	2.83	3.01	3.52
Hypothyroidism (12)	1.18	1.24	1.05	1.18	0.48	0.57	0.49	0.48
Kidney stones (7)	0 *	0 *	0	0	0.65	0.75	0.80	0.94
Peptic ulcer (23)	0.83	0.88	0.82	0.90	0.76	0.89	0.87	1.02
Low birth weight (59)	0.45		0.55		0.72		0.81	
Miscarriage (88)	0.98		0.87		0.82		0.77	
Prematurity (7)	0		0		1.23		1.39	
Birth defects (11)	0.45		0		1.61		1.50	
Hives (37)	0.56		0.67		0.36 *		0.45 *	
Skin rash (27)	1.46		1.23		0.39 *		0.35 *	
Acne (24)	0.10 *		0.12 *		0.68		0.70	
Psoriasis (13)	1.51		2.37		1.65		6.10 *	
Eczema (51)	1.24		1.29		0.93		0.97	
Specified allergies (40)	0.89		0.82		0.65		0.62	
Asthma or hay fever (49)	0.88		0.93		0.50		0.51	

Since the distribution of diseases is never uniform, ORs and RRs always differ from unity (1:1), which is the dividing line between an excess or a deficit of cases. A ratio above unity points to an excess of cases, a ratio below unity to a deficit. A low p-value indicates whether the difference from unity is large enough to be significant, but it does not address the cause of the difference.

Significant differences for ratios *below* unity were found for hypertension, kidney stones, gall bladder disease, hives, and skin rashes. Diabetes and psoriasis were the only disorders that showed a significant excess in the exposed cohort for one or two exposure expressions.

For a better evaluation of the association of exposure and disease, a detailed analysis is necessary, taking into account the dose-response relationship, the consistency of the findings across exposure expressions and populations, and the outcomes of the multivariable analysis. Based on these considerations, the results of the analysis were not considered evidence of a positive association of diabetes with exposure to VOCs. For psoriasis, the ORs for all exposure expressions were larger than unity with a consistently positive dose-response relationship, but the number of cases was too small for multivariable analysis. However, skin and allergic disorders were the only disorders for which the case diagnosis and the date of first diagnosis (TDIAG), could not be validated due to the lack of supporting medical records in the majority of cases.

Experience with other disorders showed that errors as large as years were not uncommon in interviewee-reported TDIAG. It is possible that such errors were more frequent among members of the exposed cohort as a result of so-called recall or respondent bias. Exposed people may be more concerned about the potential effects of exposure than are unexposed people, particularly if the exposure has been publicized. Therefore, they will recall more diseases when interviewed, and the TDIAG given will tend to be more recent. This effect would entirely explain the findings for psoriasis. An error of just two or three cases with regard to the eligibility status and a proper exposure-disease sequence would drastically change the association of exposure with psoriasis, because of the small number of cases.

Another likely explanation for the positive ratio is that the comparison of multiple diseases with the same exposure is bound to yield some positive associations merely by chance. The fact that only one disorder (out of the 20 listed in Table 0.3) has shown a positive association with VOC exposure is most likely an indication of such a chance effect. This agrees with the observation that, for most diseases and abnormal laboratory values, the reference population was at increased risk, and not the exposed cohort. For this reason, the results of the analysis for psoriasis were considered, at best, insufficient evidence of a toxic effect of VOCs. In summary, there was no evidence of an observable adverse effect of exposure to VOCs in drinking water at the prevailing levels of contamination. This can be translated into a NOEL (No Observed Effect Level) of:

- * A *median* available concentration of 84 parts per billion (ppb) of *total* VOCs in water (*maximum* 13455 ppb);
- * A *median* TAEVOC of 4666 ppb-months accumulated during the entire exposure period (*maximum* TAEVOC of 393 million ppb-months);
- * A *median* DOSEVOC of 15525 units, equal to an ingested dose of 3.7 mg/day total VOCs for a period of 47 months (*maximum* dose of 477 mg/day, or 6.8 mg/kg/day assuming an average weight of 70 kg, for a period of 99 months).

In this study, the *maximum* dose is approximately the same as the NOEL found for *single* VOCs in animal experiments and human occupational studies (listed in Chapter 3). The current study used a much wider array of health indicators. The inference that the lack of evidence of

an adverse effect of VOC exposure did not result from too small a study size is permitted by the fact that the maximum NOEL approximates NOELs from other human studies and animal experiments.

The conclusion of no adverse effects of VOC exposure does not alter the fact that there were differences in disease occurrence between the cohorts, as shown in Table 0.2. Further, most disorders were more frequent than expected from national data. As shown in Table 0.2, 10 of the 21 disorders were more frequent in the contaminated areas than in the reference areas. This approximately one-to-one ratio (10 more frequent to 11 less frequent) is expected if differences are determined solely by chance. The ratio is higher if only statistically significant differences are considered, however.

This raises a question about the true cause(s) of the larger differences in disease occurrence. A recall bias may explain the differences for skin rashes and psoriasis, since no validation of the self-reported diagnosis could be carried out. For the other disorders, chance may be the dominant cause, because multiple comparisons were made (the effect of which is not reflected in a simple p-value).

Another possible explanation may be that the study population may have had less access to the health care system than the population in general; also, within the study population, the Battle Creek cohort may have been in a more favorable position than the other cohorts. Differences in health care delivery from neighborhood to neighborhood, may lead to differences in disease rates. Such differences are misleading as they do not reflect the true situation, but rather the efficacy of the system in detecting and/or treating people in need of medical care. In Chapter 9 of this report, several issues are identified which warrant further exploration, since they might reveal options for improving the health care system.

This indication of potential problems in the health care system certainly does *not* imply that they are specific to the study area. They may reflect a general problem which surfaced only because of the depth and breadth of this study. Most of the prevalence and incidence rates found were higher than those estimated for the U.S.A. as a whole. A higher degree of case-finding, as accomplished in this study, in combination with a wider variety of compiled data and a more detailed analysis, may have unveiled problems which otherwise might not have surfaced.

A separate study component comprised the evaluation of mortality rates since 1970 and hospital discharge rates in 1982-1983 for small political divisions. In this analysis, the liver disease mortality rate showed a statistically significant increase, most likely related to the increased rate of alcohol-abuse morbidity. Excess cancer mortality was observed in Battle Creek, but not in Emmett and Pennfield Townships which harbor the exposed Verona Park neighborhood. The value of these findings is unclear. Cancer mortality is a poor indicator of the occurrence of cancer, because it does not include cases who survive. Hospital discharge data are, in theory, a better measure of how frequently diseases occur; at present, they are inherently inaccurate because of inability to distinguish between first and repeat admissions. Thus, patients may have been counted multiple times for the same disease. However, both information sources may have an important public health value which should be further explored.

0.11 Conclusion

The Battle Creek Health Study is a comprehensive epidemiologic study of the potential health effects of chronic exposure to low levels of VOCs in drinking water. The study was specifically designed to address the methodological problems inherent in the assessment of the health risks from simultaneous chronic exposure to a number of toxic compounds with multiple

routes of entry. The study involved 251 current and former residents of the contaminated area and 498 of the reference area. Of the 251 potentially exposed participants, 28 had left the exposed area before their well water became contaminated. The maximum estimated exposure level among the remaining 223 exposed people was 3.3 grams total accumulated VOCs in drinking water (available exposure) or 6.8 mg/kg/day (the dose, based on the actual volume of water consumed). No excess of health disorders was observed that could have been attributed to the VOCs. Alternative analyses included: the compounding effect of exposure to chloroform, a byproduct of chlorination of drinking water supplies; a study population limited to the Verona exposed and the Calhoun County reference population; and occupational exposure to VOCs (if the number of cases was sufficiently large for multivariable analysis). These alternative analyses yielded results consistent with this conclusion.

Although not attributable to VOCs, diabetes, cancer in women, miscarriages, skin rashes, and psoriasis occurred more frequently, whereas gall bladder disease, hypertension, acne, and hives occurred less frequently in the contaminated areas than in the comparison areas. Other differences did not reach statistical significance. The available information is insufficient to explain either the surplus or the deficit of disease.

There is information suggesting some important gaps in the information exchange between physicians and patients; difficulties in obtaining access to the health care system; and possibly poor therapy compliance and over-medication. The information was not specifically compiled for the study of these issues, however, and further studies have been recommended.

This study has yielded valuable information on population-based incidence rates of many diseases. Except for information on cancer, there have been no reports in the published literature with information on incidence rates of other diseases which cover all ages and both sexes, followed up for a sufficient period of time.

Chapter 1 Introduction

1.1 The Initial Situation

Contamination of groundwater supplies by toxic chemicals is considered to be one of the most important environmental health threats facing the State of Michigan. Approximately 50% of the population, about 4.5 million people, depend on groundwater as the sole source of drinking water. Michigan has inventoried approximately 1,000 known or suspected sites of groundwater contamination impacting on at least 25 community water supplies and hundreds of private wells statewide. Among the most commonly detected contaminants are the volatile organic compounds (VOCs) of concern to this study. In 1982, the U.S. Environmental Protection Agency (EPA) found that approximately 28% of community wells, from systems serving populations over 10,000 people, were contaminated with at least one VOC (24).

In August of 1981, VOCs were detected in a water sample collected from a church served by the municipal water supply of the Battle Creek City, Calhoun County, Michigan. This led to a program for monitoring VOCs in the 30 city wells. The well field, which serves a population of 38,000, is in the northeastern corner of the city, adjoining both Emmett and Pennfield townships. The nearest of the suburban residential developments surrounding the well field is Verona Park, hence the name Verona Well Field.

The households in the area of the Verona Park Well Field obtained their drinking water from private wells. In November of 1981, some residential wells in Verona Park were also found to be contaminated, and further testing revealed more affected wells. At the request of the Michigan Department of Natural Resources (MDNR), the EPA studied the extent of the aquifer contamination, and confirmed the presence of a plume of VOCs traveling in a northwesterly direction toward the Battle Creek river, traversing Verona Park and the city's well field. The source of the contamination appeared to be leaking storage tanks at a local solvent reclaiming and distributing company. In February, 1984, free floating VOCs were detected on top of the water table underneath the solvent company, while VOCs at levels up to several hundred parts per million (ppm) were found in soil samples. A possible second source of contamination was the nearby railroad staging yard, where VOCs were suspected of being either spilled or used for degreasing the equipment.

The Center for Environmental Health and Injury Control (CEHIC, then known as the Center for Environmental Health), of the Centers for Disease Control (CDC), judged that the VOC levels observed in the wells might pose a potential health threat. A recommendation was made to close city wells with VOC concentrations exceeding the levels associated with the upper bound of the estimated lifetime excess cancer risk of 1 in 100,000. In June of 1982, the Michigan Department of Public Health (MDPH) initiated a program to provide bottled water to households with contaminated wells, and in June of 1983, the MDNR requested the EPA to conduct a remedial investigation and feasibility study under the Comprehensive Environmental Response, Compensation and Liability Act.

In August of 1983, a project was started to provide for the extension of municipal water mains in the Verona Park area. Reduction of the VOC concentrations in the city's drinking water was accomplished by using only those wells known to have low concentrations in an on-line sequence, as defined by the MDPH based on the water demand of the system. By these measures, the VOC levels in the blended city water leaving the water plant was kept at 8 parts per billion (ppb) or less of total VOCs. In 1984, further improvement was achieved when new and VOC-free city wells came into operation. Since then, the city's tap water has been either free of VOCs, or the total VOC concentration has not exceeded 2 ppb.

In 1983, similar VOCs were detected in residential well water in Dowagiac, Cass County, Michigan, about 70 miles southwest of Battle Creek, and in the city of Springfield, southwest of and adjacent to Battle Creek. In Dowagiac, the source of contamination appeared to be an industrial plant that used VOCs for degreasing manufactured metal parts. In Springfield, no source was found.

Following a request from concerned citizens and the health department of Calhoun County, the MDPH felt that an epidemiologic study was indicated, encompassing the three residential areas with contaminated private wells. The CEHIC agreed that a study could be meaningful for two reasons: 1) the scientific prospects of the study, and 2) its value to public health and public health policy. At the prevailing VOC levels, toxic effects were unlikely to be observed (Chapter 3); this expectation was based on extrapolating effects observed at high dose levels to the ultralow exposure levels of environmental contamination. However, it is within the scope of public health policy to deal with the uncertainties inherent in this extrapolation. Low dose studies have to be conducted to provide a data base of scientific facts for dealing with public concerns. Even a negative study could still contribute valuable information. The results of several such studies, each by itself too small to be conclusive, might be pooled to form one data base, large enough to yield useful conclusions pertaining to the health effects of low dose VOC exposure.

1.2 The Contaminants

Volatile organic compounds or VOCs are distinguished from other volatiles by certain physico-chemical properties of importance in testing chemicals. VOCs include aliphatics as well as simple aromatics, many of which are halogenated.

In this report, however, the term VOC is used exclusively to identify, as a group or as single chemicals, the seven chlorinated C2-aliphatics, found in the study areas and shown in Figure 1.1 below.

With the exception of PCE, the volatility of these VOCs is reflected in their low boiling points of 32 - 87°C. Although PCE's boiling point of 121 °C is much higher than that of water, it is still considered a volatile. VOCs are used in the chemical industry as solvents and as intermediates for a large number of consumer and technical products, such as upholstery, food wrappings, coatings for tanks and pipes, paper coatings, house paints, brush cleaners, vinyl plastics, additives to leaded gasoline (12DCA is a lead-scavenger), etc. Based on their excellent capacity to dissolve fats and oils, VOCs are used to extract edible oils, and to degrease manufactured metal parts and electrical components. VOCs have been used for decades, some for more than a century, to induce local and general anesthesia. In the USA, they have now been replaced with safer products. To prevent or retard chemical degradation, stabilizers are added to VOCs. It is worth mentioning, however, that these additives may not totally prevent the formation of hydrochloric acid, a major degradation product of VOCs, and probably the main cause of leaking tanks.

The seven VOCs have a distinctive, somewhat sweet odor, but the odor threshold varies considerably among individuals, ranging from 50 to 300 parts per million (ppm) in air, with outliers as low as 3 and as high as 700 ppm (27, 29, 32, 34). The odor threshold for VOCs in water depends on the degree to which the chemical is released from the water into the air. This is known as the water/air partition coefficient. The coefficient is determined by the chemical's solubility in water, and the temperature of water and air. Odor thresholds of 300 and 500 ppb in water have been reported for PCE and TCE respectively (24, 32). Thus, odor is an insensitive measure of detection.

Name	Abbreviation	Structure
1,1-dichloroethane	11DCA	$\begin{array}{c} \text{Cl} \quad \text{H} \\ \quad \\ \text{Cl}-\text{C}-\text{C}-\text{H} \\ \quad \\ \text{H} \quad \text{H} \end{array}$
1,2-dichloroethane or ethylene dichloride	12DCA	$\begin{array}{c} \text{H} \quad \text{H} \\ \quad \\ \text{Cl}-\text{C}-\text{C}-\text{Cl} \\ \quad \\ \text{H} \quad \text{H} \end{array}$
1,1,1-trichloroethane	TCA	$\begin{array}{c} \text{Cl} \quad \text{H} \\ \quad \\ \text{Cl}-\text{C}-\text{C}-\text{H} \\ \quad \\ \text{Cl} \quad \text{H} \end{array}$
1,1-dichloroethylene or vinylidene chloride	DCE	$\begin{array}{c} \text{Cl} \quad \text{H} \\ \quad \\ \text{Cl}-\text{C}=\text{C}-\text{H} \end{array}$
cis-1,2-dichloroethylene	CIS	$\begin{array}{c} \text{Cl} \quad \text{Cl} \\ \quad \\ \text{H}-\text{C}=\text{C}-\text{H} \end{array}$
1,1,2-trichloroethylene	TCE	$\begin{array}{c} \text{Cl} \quad \text{H} \\ \quad \\ \text{Cl}-\text{C}=\text{C}-\text{Cl} \end{array}$
1,1,2,2-tetrachloroethylene or perchloroethylene	PCE	$\begin{array}{c} \text{Cl} \quad \text{Cl} \\ \quad \\ \text{Cl}-\text{C}=\text{C}-\text{Cl} \end{array}$

Figure 1.1 Chemical names and structures of the seven chlorinated C2-aliphatics found in the study area and to which the term VOC is exclusively applied in this report.

1.3 The Sites

Figure 1.2 depicts the location of city and private wells in the Verona area, and of the source of pollution. The Springfield site can be characterized as a small number of residences with contaminated wells, the locations of which have no specific pattern suggestive of a common source. In Dowagiac, the polluted wells form a small compact area adjacent to a manufacturing plant storing large volumes of solvents in underground tanks. The contamination level in private wells in the study area varied widely within and among neighborhoods.

Table 1.1 gives an overview of the concentrations observed in city and private wells. A test result of zero ppb indicates absence of the chemical, or its presence at a level too low to be quantified. Table 1.1 shows wide differences in the contamination level within and among the areas. In Verona Park, concentrations were highest in the residential area west of the source (Area 2 in Figure 1.2), and lowest in the areas north and south of the source (Areas 3 and 4), which is in agreement with area locations relative to the main direction of groundwater flow. Trans-1,2-dichloroethylene (TRANS), a stereo-isomer of CIS, was found only incidentally in city wells, and only since 1984. It was less uncommon, though still rare, in private wells. Due to its rareness, TRANS could not be included in the model developed for transport of VOCs in the groundwater, described in Chapter 4 and Appendix C, and was ignored in the estimation of individual exposures.

RESIDENTIAL WELL SAMPLE LOCATION MAP

HAS BEEN REDACTED – ONE PAGE

CONTAINS POTENTIAL PERSONALLY-IDENTIFYING INFORMATION

Table 1.1 Summary of VOC concentrations (ppb). Wells not contaminated with at least one of the indicated chemicals have not been listed. Medians were calculated from positive tests.

	TCA	11DCA	12DCA	DCE	CIS	TCE	PCE	TRANS	Any VOC
Battle Creek city wells *									
Number of wells polluted	22	19	4	13	20	16	17	?	27
Maximum concentration	180	43	17	11	229	67	120	9	
Median concentration	5	5	2	2	5	4	5	**	
Verona residential wells West									
Number of wells polluted	36	43	38	31	46	32	24	16	58
Maximum concentration	134	150	325	82	3900	657	290	31	
Median concentration	6	18	20	4	293	47	44	5	
Verona residential wells North ***									
Number of wells polluted	3	0	0	0	0	0	3	0	3
Maximum concentration	3						1		
Verona residential wells South ***									
Number of wells polluted	1	2	4	1	11	5	5	0	13
Maximum concentration	1	1	2	3	9	14	9		
Dowagiac residential wells									
Number of wells polluted	10	8	1	7	15	15	7	6	17
Maximum concentration	935	60	1	90	1919	14890	8	5	
Median concentration	309	16	1	31	226	3214	4	3	
Springfield residential wells									
Number of wells polluted	9	11	1	2	11	4	0	2	13
Maximum concentration	125	547	10	96	81	6	-	1	
Median concentration	6	4	10	51	16	5	-	1	

* The concentration values of the city wells are derived from the monitoring program and, thus, refer to many samples per well. In contrast, private wells typically had only one sample; a few had two.

** In city wells, TRANS was a very rare finding. Values were usually not reported, and a median was therefore not calculated.

*** A median was not calculated, as only a few wells were polluted, and the concentrations were quite low (close to 1 ppb).

1.4 Phase I

In February of 1984, a memorandum of understanding was signed between the CEHIC and the MDPH, outlining the interests and responsibilities of both agencies in a health study. A public meeting was held in Battle Creek on March 15, 1984, to present the essentials of a two-tiered health study to residents of the areas surrounding the Verona Well Field. Phase I was intended to gather demographic data to be used for estimating the size of the exposed population, identifying comparison areas, and for establishing selection criteria for participants in the analytical Phase II of the study. In March of 1984, a contract for Phase I, funded by the State of Michigan (\$23,879), was signed by the MDPH and the Kercher Center for Social Research, Western Michigan University in Kalamazoo, Michigan. Phase I had the following objectives:

- 1) To conduct a telephone interview with one person per household. The survey should include all dwellings with a contaminated well and all dwellings in the designated reference areas.
- 2) To provide sufficient information to characterize households in terms of size, age and sex distribution, socio-economic class, and the dates of moving into or leaving the residence.
- 3) To provide an inventory of health concerns which could be focused upon in the analytical Phase II of the study.
- 4) To provide an estimate of the size of the VOC-exposed population, and of the number of exposed and unexposed people willing to participate in Phase II of the study.

1.5 Phase II

In February of 1985, a Cooperative Agreement (# U61/CCU500838-01) was signed between the MDPH and the CDC to initiate and complete Phase II. The agreement was in effect from March 1, 1985 to February 28, 1986. Its purpose was for the CDC to assist the MDPH in designing, conducting, and analyzing an epidemiologic study of the potential health effects resulting from chronic exposure to low levels of VOCs in drinking water supplies. More specific study objectives were:

- 1) To investigate whether there is an excess of kidney and liver disease in the exposed population of Calhoun and Cass counties.
- 2) To investigate whether there is an excess cancer incidence or mortality in the exposed study population of Calhoun and Cass counties.
- 3) To investigate whether a fetotoxic effect of VOCs has occurred.
- 4) To investigate whether there is, in general, an excess of diseases other than those mentioned above.
- 5) To evaluate, in addition, observed excesses of disease in the context of county and state-wide hospital and mortality data.

A protocol for the study was completed by August 1984 and submitted for funding to the EPA through the Agency for Toxic Substances and Disease Registry. An Interagency Agreement between the EPA and the CDC provided the channel for funding the major part of Phase II. The EPA granted the MDPH \$826,827 to conduct and complete Phase II of the Battle Creek Health Study. A Request For Proposal for the field part of the study, submitted by the MDPH

for open bidding, yielded only one bidder. A contract was signed between the MDPH and the bidder, the Kercher Center for Social Research, Western Michigan University, Kalamazoo, Michigan. In the remainder of this report, the Kercher Center is referred to as the contractor. The duration of the contract, from May 31, 1985 to February 28, 1986, was extended until August 31, 1986 without additional cost in order to complete editing and computerization of the compiled data. The MDPH contracted with the Calhoun and Cass County Health Departments for collection of water samples from specified private wells within their respective counties.

The complexity of the study design required a careful consideration of the available options, good planning, and a proper distribution of responsibilities for the various study components. As mentioned earlier, the MDPH and the CEHC were the lead agencies. Only two months after publishing a Request for Proposal from potential study contractors, and one month after the contract for the field part of the study was signed, the field component was in full operation. Except for the computerization of the compiled data, the field study was completed according to time schedule and budget. Computerization of the data took half a year more, due to the vast amount of data. The large volume of data compiled - and the resulting large numbers of potential inconsistencies and errors - were clearly underestimated from the outset.

1.6 Persons and Institutions Involved in the Battle Creek Health Study, 1985 - 1987

The major parties in the study organization were:

- The MDPH: Its Center for Environmental Health Sciences had lead responsibility for the conduct of the study. The Bureau of Laboratory and Epidemiological Services did the water testing and blood serum testing for exogenous chemicals. The MDPH Center for Health Statistics was responsible for providing morbidity and mortality rates for minor civil divisions, counties, and the State.

MDPH: Center for Environmental Health Sciences

Arthur W. Bloomer, MS, Project Officer for the State
Adrian J. Oudbier, MPH, Manager, field operations
Robert L. Welch, MS, chlorinated pesticides, PCB, PBB

MDPH: Bureau of Laboratory and Epidemiological Services

Theodore J. Williams, PhD, testing of water samples

MDPH: Center for Health Statistics

Janet T. Eyster, PhD and Maurice C. Barone, state morbidity/mortality data

- The CDC: Its Center for Environmental Health and Injury Control had lead responsibility for the research design, the coordination of the various study components, and study analysis.

CDC: Center for Environmental Health and Injury Control

Stan C. Freni, MD, PhD, DrPH, Principal Investigator and Project Officer for the CDC
Donald L. Phillips, PhD, laboratory quality control, assistance in groundwater modeling
Debbi L. Kotlovker, assistance in data transfer
Paul M. Ridker, MD, assistance in health data validation

- * **Western Michigan University:** Its Kercher Center for Social Research was contracted to do interviewing, tracking, clinical examination, and abstracts of medical records. The latter two were subcontracted to Borgess Medical Center, Kalamazoo, Michigan.

WMU: Kercher Center for Social Research (Contractor for Health Study)

James C. Petersen, PhD, Contractor, health survey
Janet L. van Valey, MA, Manager, field operations
Thomas L. van Valey, Ph.D., Supervisor, data management

Borgess Medical Center (Subcontractor for clinical chemistry, clinics, and medical records)

Georgiann Ellis, Project manager
Sharlene K. Dolman, RRA, medical records
Kathy L. Koets, clinics

The successful conduct and completion of the field component of Phase I and Phase II of the Battle Creek study could not have been accomplished without a thoughtful and anticipatory organization. In turn, no organization could have achieved this outcome without the greatly appreciated support of the various components and individuals within the MDPH and the CDC, the cooperation of hospital administrators and private physicians who made their medical files available, and the cooperation of the residents involved. The support received from the MDNR, the Calhoun and Cass County Health Departments, the Governor's office, the regional EPA, and the US Geological Survey was also greatly appreciated. They provided information and assistance pertaining to water quality, hydrogeologic investigations, and residential data, or were the liaison in contacts with other government offices and the public. Without diminishing in any way the value of the contribution of others, the authors believe that special acknowledgement should be made of the help received from Mrs. Janet L. van Valey and Mr. Adrian J. Oudbier. As manager of field operations (Mrs. van Valey for the contractor, Mr. Oudbier for the MDPH), both were instrumental in the timely completion of the data gathering components of the study. The organization underlying this complex, yet smoothly and effectively running operation, is depicted in the listing above.

1.7 Contents of the Report

The final analysis of the compiled data did not start until September 1986, when the questionnaire and laboratory data and the medical records abstracts were submitted to the CEHIC by the contractor. The wealth of information obtained in this study provided opportunities for analyzing health events regarding issues other than exposure to VOCs. Such an analysis may take several years for completion. The approaching expiration of the Interagency Agreement of the CDC and the EPA called for an earlier report, however. This Report addresses the primary objective of Phase II of the Battle Creek Health Study, which is essentially: "Investigate whether there is an observable health effect of exposure to VOCs in drinking water".

Chapter 2 presents a summary of the methodology. More details are provided in Chapters 3 through 7. They describe the major steps in risk assessment: hazard identification or toxicology of VOCs (Chapter 3), the methodology specifically developed for exposure assessment (Chapter 4), and the assessment of health status (Chapter 5). Chapter 6 gives an outline of the concepts involved in the epidemiologic evaluation and statistical analysis of potential associations between exposure to contaminated drinking water and health effects. The quality control and assurance procedures for laboratory tests are described in Chapter 7. Quality issues regarding data for factors other than exposure have been addressed in paragraphs interspersed throughout Chapters

2 through 5. The results of the study are reported in Chapter 8 and in several appendices. An extensive discussion of difficult issues is given in Chapter 9. This chapter also addresses the final step in risk assessment, that is, the characterization of the risk estimates, and how these relate to the uncertainties inherent in the data base or in the methods applied. The Report contains appendices with a detailed description of a component of the exposure assessment (groundwater modeling) and some disease-specific analyses and results. An Executive Summary provides a comprehensive summary of the study and its conclusions.

Much of the methodology is novel, developed specifically for this study, but with an intended general applicability to other scenarios. This required extensive discussion in the relevant chapters, which to some extent explains the large volume of this report. The authors anticipated that many readers would be unfamiliar with toxicological, epidemiological, and risk assessment issues, and desired to provide as much basic data as possible to enable the informed reader to reach conclusions independent of those presented by the authors. The details and explanations provided, and the attempt to avoid professional jargon as much as possible, have resulted in a voluminous report. Careful reading is necessary to avoid confusion. Therefore, rather than reserving all discussions for one chapter, discussion paragraphs have been inserted in the relevant chapters and sections.

The presentation of the cost of the various field elements of the study is unusual in epidemiological reporting. The authors appreciate the contractor's willingness to provide information on cost components and procedural issues, which is attached as Appendix B. This cost information will prove invaluable in the design of future studies. It would have helped the authors to improve this study, while controlling the cost, if this type of information had been available during the design and conduct of the Battle Creek Health Study.

1.8 Peer Review

Many of the analytical and methodological issues of the Battle Creek Health Study have been discussed with scientists not involved in the study itself. Especially acknowledged is the help received from: Dana Flanders, MD, DrPH; Phillip Rhodes, MS; and Riduan Joesoef, MD, PhD (all from the CDC); Robert Abbott, PhD (from the National Institute of Health); and Norman Granneman, MS (from the U.S. Geological Service in Michigan). In addition, complying with regulations, this Report was formally reviewed by a panel of outside experts consisting of:

**John K. Hawley, PhD, Director, Bureau of Toxic Substance Assessment,
New York State Department of Health, Albany, New York;**

**Carl M. Shy, MD, MPH, Professor of Epidemiology, School of Public Health,
University of North Carolina, Chapel Hill, North Carolina;**

**Daniel F. Smith, DrPH (replacing Raymond R. Neutra, MD, DrPH),
California State Department of Health Services, Berkeley, California;**

**Richard Wilson, PhD, Professor of Physics, Energy & Environmental Policy
Center, Harvard University, Cambridge, Massachusetts.**

The Panel summarized its comments in writing and verbally at a March 21, 1988 meeting at the CDC in Atlanta, Georgia, chaired by Daniel A. Hoffman, PhD, Assistant Director for Science, Division of Environmental Hazards and Health Evaluation, CEHIC, CDC. The comments generally favored the scientific contents and conclusions of the report, although critical comments were received on the presentation of methods and results. This version of the report is revised in agreement with these comments.

Chapter 2 Summary of the Methodology

2.1 Introduction

In the Battle Creek Study, the exposure factor is predefined, and the search is for an association with diseases which may follow exposure. During a defined period of follow-up, a group of people exposed in the past (exposed cohort) is compared to a group of unexposed people (reference or comparison cohort) with regard to their past health experience. If potential confounding factors are adequately controlled, significant differences in disease incidence between cohorts may then suggest that the disease is due to, or at least associated with, the chemical(s) of interest. It is essential to recognize, however, that the exposure levels are low, far below levels which produced no adverse effects in animal or occupational studies (Chapter 3). Thus, low exposure levels, combined with the rareness of most diseases, render statistical exposure-effect associations, if found, more likely to be due to chance than to a toxic effect of the exposure. To reduce the chance of spurious associations, emphasis was given to a detailed exposure assessment and to a quality control program for laboratory tests as well as for health data.

An epidemiologic study of the effects of toxic chemicals will yield optimal information about the nature of observed statistical associations, if its design adheres to risk assessment principles:

- * The study must address the health risk from single chemicals and from a defined chemical mixture.
- * The exposure status must be accurately assessed for each individual separately, while focusing on dose rather than the available amount of chemicals. The study must account for secondary routes of entry and exposure.
- * Quality control procedures must apply to both exposure and health data.
- * Exposure incurred after a disease is diagnosed, as well as any disease diagnosed prior to exposure, must be discounted prior to estimating the risk.
- * The risk must be evaluated in the context of a dose-response analysis.
- * The analysis must properly control for the effect of confounding factors (to compensate for the fact that people cannot be experimentally subjected to exposure).
- * The effect of uncertainties in data and methods on the outcome of the analysis must be evaluated.

2.2 Statistical Power as a Basis for the Study Size

The statistical power of a study (the probability of finding a health effect of exposure, if it truly exists) is determined by the numbers of participants in the exposed and unexposed cohorts, the ratio of these numbers, the background rate of occurrence (the incidence) of the disease of interest, the quantitative potency of the chemical involved (excess risk of disease per unit of exposure), and the threshold of statistical significance (alpha) that one would accept (usually 0.1 or 0.05). The only disease for which there is information on the background incidence and the potency of VOCs is cancer. However, VOCs have not been shown to be carcinogenic to humans, and the extrapolation of human risk from animal data is complicated by the current lack of understanding of the mechanisms of carcinogenesis. Public health authorities have dealt with this problem by making conservative assumptions (overestimating human risk)

when extrapolating risk from animal studies to humans and from high-dose observations to low-dose expectations. It is estimated that 2 to 18 ppb of PCE, TCE, and 1,2-DCA, and 220 ppb of TCA may be the lowest VOC concentrations in drinking water associated with a lifetime cancer risk of 1 per 100,000 at a consumption of 2 liters of water a day (24). Even if the entire exposed cohort had a lifetime exposure to a 1000 times higher contamination level (the real exposure was for less than 15 years), the power of a study of 250 exposed and 500 unexposed persons would still not exceed the lower power limit set by the alpha-level. For an alpha of 0.05, the study would have a chance of not more than 1 in 20 of revealing an excess cancer incidence, if a cancer risk exists. This renders cancer risk meaningless for power calculation.

Power calculations may be based on the incidence rate of diseases other than cancer. However, national surveys have shown that the prevalence of most chronic diseases is lower than or about the same as that of all cancer combined (69), which suggests that for these diseases the statistical power of the study is not markedly different from that for cancer. There are no dose-response data for VOCs and chronic-diseases-other-than-cancer in either animals or humans. The conventional way of computing the power for a given relative risk, say 1.5 or 2, has no meaning in this case, since it does not take into account the potential of the chemical to cause an effect. Rather than base the study size on power calculations, it was decided that Phase II should encompass the entire exposed residential population, and twice that number from reference neighborhoods. A ratio larger than 2:1 would not significantly have increased the probability of success, and would certainly not be justified from a cost-effectiveness perspective.

2.3 Phase I

2.3.1 Selection of Participants

Phase I consisted of selecting and surveying households for the two study categories: "exposed" and "reference" or "comparison" (unexposed, but otherwise comparable households). Reference neighborhoods were selected for their similarity to the exposed areas, based upon gross estimation of the real estate value, size, and age of the dwellings, and the presence of wells (neighborhoods in Battle Creek City had no private wells). To enhance the homogeneity of the study population, the search was for complete blocks or entire communities rather than individual houses. Although Phase II called for a 2:1 ratio, Phase I comparison neighborhoods were chosen to be large enough to yield four to five reference households for each exposed one. Disparities in household characteristics and refusals to participate in Phase II were expected to substantially reduce the pool of reference households ultimately available for Phase II (the actual epidemiological study). The neighborhoods selected for Phase I of the study were:

- * All households with contaminated private wells in Verona Park, Dowagiac, and Springfield (the exposed cohort).
- * All households in selected blocks in the city of Battle Creek City (BC city cohort). These blocks were connected to city water.
- * All households in three blocks of Brownlee Park southeast of the Verona Well Field, and the entire population of Ceresco, a small community 8 miles east of Battle Creek City, both with uncontaminated private wells as the source of drinking water. These neighborhoods and the BC city group formed the Calhoun County reference cohort for the Verona exposed cohort.
- * All households near Edwardsburg and near Barron Lake, both in Cass county, with uncontaminated private wells. This group formed the Cass County reference cohort for the Dowagiac exposed cohort.

"Exposed", in this context, means that the members of the cohort had at some time lived in a dwelling with a well found contaminated in the 1981-1983 sampling survey. Whether or not these individuals actually had been exposed to VOCs was to be established in the individual exposure assessment as part of Phase II (see below).

2.3.2 Data Gathering - Interviews

In Phase I, the interviewee was usually the head of the household or his/her spouse. Data were compiled on household size, income, sex and age composition, the date the respondent moved into the dwelling, education and occupation of the two main adults, whether the respondent or any other adult in the household was occupationally exposed to chemicals, the water source, the date the interviewee thought that pollution of the water might have started, willingness to participate in the planned Phase II of the study (the respondent was assumed to represent all members of the household), and an open-ended question about the interviewee's health concerns. The income level was the gross family income per annum. The descriptors for coding of the responses were:

School level	Income level	Occupational exposure
1 = grade school	1 = under \$10,000	0 = no one exposed
2 = some high school	2 = \$10 - 20,000	1 = at least one exposed
3 = high school completed	3 = \$20 - 30,000	
4 = some college	4 = \$30 - 40,000	
5 = college completed	5 = \$40 - 50,000	
6 = higher education	6 = over \$50,000	

2.3.3 Demographic Descriptors

Tables 2.1 and 2.2 show the results of Phase I in terms of demographic descriptors and willingness to participate in Phase II. As expected, reference households were less inclined to participate. Table 2.2 shows only a few small but statistically significant differences in demographic characteristics between households willing to participate in Phase II and "others", i.e., households who refused to participate, or were not yet sure. Households indicating a willingness to participate in Phase II were larger in size and had a lower percentage of adults (age 20+).

If the households are grouped according to exposure status, the differences between populations are smaller still. An unexpected finding, of importance in the analysis of Phase II, was that exposed households had a *lower* mean percentage of people occupationally exposed to chemicals, and that households intending to participate in Phase II had a *higher* mean percentage of such persons. The annual income of exposed households was, on average, about \$4000 less than that of reference households. Respondents often refused to provide information on the household income, and differences in Tables 2.1 and 2.2 may therefore not reflect the real situation. However, the lower income level is compatible with the slightly lower education level. In summary, although reference neighborhoods were selected on the basis of "on sight"

comparability, the overall similarity in the demographic and occupational parameters of the Phase I population was quite satisfactory.

Table 2.1 Demographic descriptors (mean and %) of Phase I households.

Household location & type (N)	year to area [#]	househ size	age	% adult	% male	school level	income level	% occup exposure
Verona exposed								
participate (52)	1968	3.0	31.3*	62	46	1.9	1.6	23
all other (9)	1966	2.8	39.3	72	40	2.0	1.9	22
Springfield exposed								
participate (5)	1964	3.0	31.5*	60	40	2.3	1.3	60
all other (3)	1956	1.7	56.0	80	80		2.0	33
BC city reference								
participate (118)	1970*	2.7	31.8*	65*	47	2.5*	1.7	42
all other (61)	1965	2.5	37.3	73	52	2.3	1.6	33
Brownlee reference								
participate (60)	1969	2.7	32.9*	71	54	2.2	2.3	50*
all other (48)	1967	2.5	37.8	73	53	2.1	2.2	31
Ceresco reference								
participate (30)	1965	3.1	33.4*	65*	50	2.5	2.8*	73
all other (19)	1958	2.3	46.2	81	44	2.6	2.2	53
Dowagiac exposed								
participate (14)	1964	3.4	33.7*	64	57	2.6	2.6	64
all other (2)	1967	2.0	64.5	100	25	2.0	1.5	50
Edwardsburg reference								
participate (45)	1970*	2.7	36.8	75	48	2.5*	3.0*	44
all other (23)	1975	2.5	38.5	74	50	2.3	2.5	39
Barron Lake reference								
participate (28)	1966*	3.4*	32.3*	64*	46	2.3	2.2	54
all other (22)	1971	2.7	39.3	75	53	2.4	2.5	45

[#] "Year to area" is the year that the *respondent* moved to the study area (other household members may have different dates). School, income, and occupational exposure to chemicals refer to the two main adults in the household. All other columns refer to the entire household.

* Indicates that the difference between households intending and not intending to participate in Phase II, and between the exposed and reference households is statistically significant at p-values ranging from 0.09 to less than 0.0001 (one-sided t or χ^2 test).

Table 2.2 Demographic descriptors (mean and %) of the Phase I population by the household's intention to participate (partcpt) or not ("other" also means "don't know") in Phase II. A household is "exposed" if it is in the contaminated area; "refer" means reference household.

	Exposed population		Refer population		Total population		Total populatn	
	partcpt	other	partcpt	other	partcpt	other	exposed	refer
Number of households	71	14	281	173	352	187	85	454
Year respondent moved to area	1971	1964	1969	1967	1969	1966	1967	1968
Household size	3.1	2.4	2.8*	2.5	2.9*	2.5	3.0*	2.7
Age	31.8*	44.7	33.0*	38.8	32.8*	39.2	33.6	35.0
Percent age 20+	62*	77	67*	74	66*	74	64*	70
Percent males	48	44	47*	51	47	50	47	49
School level	2.1	2.0	2.4	2.3	2.3	2.3	2.1*	2.4
Income level	1.8	1.8	2.2	2.0	2.1	2.0	1.8	2.1
% Occup exposure	34	29	49*	37	46*	36	33*	44

See the notes to Table 2.1. In Table 2.2, * indicates a p-value ranging from 0.08 to 0.001 (one-sided t or χ^2 test).

2.4 Phase II

2.4.1 Selection of Participants

In comparing the health and exposure status of populations, attention must be given to sex and age composition, since the likelihood of developing a disease, and the degree of exposure may differ with age and sex. Potential confounding factors are the household income level (which may determine dietary quality and the degree and ease of access to adequate medical care) and VOC exposure in the workplace (occupational VOC exposure is not uncommon). Other confounding factors are usually disease-specific, and can be dealt with in the analysis.

Selection of participants for Phase II was based on the findings in Phase I. A Phase I household was considered eligible for participation in Phase II if the respondent had at least 1 year of residence in the study area prior to the intended start of Phase II. In practice, respondents who moved to the area later than 1983 were not considered for Phase II. For the exposed cohort, an additional eligibility criterion was that the well of the dwelling was proven to be contaminated with at least one of the VOCs.

Reference households were selected using the following criteria to match them to "exposed" households:

1. The year the *respondent* of the Phase I household moved to the area; the difference between exposed and reference households should be 5 years or less.
2. The household size should not differ from that of the "exposed" household by more than two persons.
3. The number of males and the number of females should not differ from the numbers of males and females in the "exposed" household by more than one person.
4. The mean age of the household should not differ from the mean age of the "exposed" household by more than 5 years.

In developing these criteria, it was expected that other descriptors would show a comparable closeness, as the neighborhoods were selected *a priori* for gross similarity. Although Phase I encompassed five times more reference households than exposed households, this number was still too small to permit adherence to the above criteria. Matching criteria were, if necessary, relaxed in a sequence the reverse of the one above. Thus, if no matching household was found, a household with a difference in mean age of up to 10 years was accepted. If still no matching household was found, a difference in the numbers of males and females of two persons was accepted, etc. In Cass County, this relaxation of the matching criteria was still too stringent to yield the requisite number of control households; this ultimately led to the acceptance of reference households smaller than the exposed household by more than two persons.

Former occupants of the dwellings in the exposed areas, and those in reference areas selected *after* the above matching procedure, were included in the study to maximize the size of the exposed cohort, and to avoid the possibility of a sampling bias, since former residents might have had a health experience different from that of current residents. A tentative analysis of groundwater data showed that the contamination of wells in the Verona area started later than 1975; therefore, tracking of former residents was limited to people who left the area in 1975 or later. To trace past residents, information on their names and current addresses was obtained from current and neighboring residents of the selected dwellings in the study area, family living in the area, city and telephone directories, and utility bills. In addition, a list of dwelling addresses and the names of past residents were published in the local newspapers, asking for information on their current address. Since selection of matched households preceded these tracking activities, *matching criteria were not applied to tracked households*. It was expected that the tracked household would have the same characteristics, on average, as the current household.

Regardless of the Phase I response with regard to participation in the next phase of the study, all addresses in the exposed areas were visited again for a Phase II interview. Some people who initially refused, or were reluctant to join the study, became participants at this point. In part, this may be attributable to the offered "incentive" of \$50 per person, a compensation for the considerable cooperation requested. Table 2.3 shows demographic parameters of the cohorts eventually canvassed for Phase II. In comparing Tables 2.1 and 2.2 with Table 2.3, it should be kept in mind that the former refer to households, not to individuals. The year a Phase I respondent moved to the area was *assumed* to be the same for the other household members; in Phase II, individual dates were known. Phase I data on occupational exposure referred to the two main adults in the household; Phase II data covered all members of the household and provided more information on the kind of chemical exposure. This facilitated the separation of people who were likely to be exposed to VOCs ("occupational exposure to VOCs" in Table 2.3). No information on education and income level was requested in Phase II. Many of

the Phase I respondents refused to respond to questions regarding their income, while the difference in the educational level was negligibly small.

Another difference between Table 2.3 on the one hand and Tables 2.1 and 2.2 on the other is that 35% of the Phase II population were former (tracked) residents, who were covered neither by the Phase I interview nor by the selection criteria for the Phase II cohorts. Tracked households were expected to resemble current households as to matching criteria, with the exception of age and when they moved to the study area. Tracked individuals were younger (by 6-8 years) and more mobile than current residents. There were no other statistically significant differences between tracked and current residents. Tracking of individuals went across the boundaries of the Phase I household. Family size and mean age of the tracked households have no bearing on Tables 2.1 and 2.2 and are, therefore, not presented in Table 2.3. Table 2.3 confirms the expectation that the tracked population would closely resemble the current population with regard to demographic characteristics.

Table 2.3 Demographic descriptors (means) of Phase II population. "Tracked" = former residents.

	Exposed population		Refer population		Total population		Total population	
	tracked	current	tracked	current	tracked	current	tracked	exposed reference
Individual data:								
Number of people	97	154	162	336	259	490	251	498
Year participant moved to area	1974*	1971	1973*	1970	1973*	1970	1972*	1971
Age	26.9*	33.7	28.2*	36.0	27.7*	35.3	31.3	33.4
Percent age 20+	68	65	66	68	67	67	66	67
Percent males	44	49	49	45	47	46	47	47
% Occup exposure	22	25	33	28	29	27	24	30
% Occup exposure to VOCs	11	12	12	13	12	13	12	13
Household data:								
Number of households	**		**		**		78	157
Household size	**		**		**		3.2	3.2

* The difference is significant at a p-value ranging from 0.07 to less than 0.0001 (one-sided t or chi² test).

** Not computed; tracking was done across household boundaries (text). A tracked household may differ in size from the initial household if some members refused to participate, or if the initial household changed because of death, divorce, births, marriage, etc.

2.4.2 Participation Rate

A conventional parameter of effectiveness in covering the target population is the refusal rate. This cannot be extracted from Tables 2.1 or 2.2, as "intention to participate" is not the same as actually taking part in Phase II, and respondents may not have been representative of the household. Tables 2.4 and 2.5 show the breakdown of the number of eligible and participating households, by neighborhood.

Table 2.4 Number of exposed households eligible for and participating in Phase II.

Exposed neighborhoods	Verona	Springfield	Dowagiac
Dwellings with well showing at least 1 VOC	82	18	17
Dwellings found uninhabited	0	-2	0
No Phase I visit due to error in address	-15	0	-1
Non-residential addresses	-6	-6	0
Refused to respond in Phase I	0	-2	0
Total: Number of exposed households in Phase I	61	8	16
Not contacted in Phase II by mistake	-1	0	0
Newly added (correction of wrong address)	6	0	0
Ineligible (wells proved Not polluted)	-2	0	0
Dwelling was uninhabited in Phase II	-3	-2	0
Refused to respond in Phase II	-11	-1	-3
Total: Phase I addresses left in Phase II	50	5	13
New households added to Phase II by tracking	7	1	2
Total: Number of households in Phase II	57	6	15
Percent refusals by neighborhood	$11/(50+11) = 18\%$	$1/(5+1) = 17\%$	$3/(13+3) = 19\%$
Percent refusals in exposed cohort	$(11+1+3)/(61+6+16) = 18\%$		

Table 2.5 Number of reference households eligible for and participating in Phase II.

Reference neighborhoods	BC city	Brownlee	Ceresco	Edwardsb	Barron L
Addresses visited in Phase I	179	108	49	68	50
Dwelling found uninhabited	-3	-1	0	1	-2
Superfluous - not selected for Phase II	-90	-72	-31	-39	-30
Total: Phase I addresses left in Phase II	86	35	18	28	18
<i>Refused to respond in Phase II</i>	-25	-8	-3	-10	-7
New households from tracking	21	1	0	3	0
Total: Number of households in Phase II	82	28	15	21	11
Proportion refusals	25/86	8/35	3/18	10/28	7/18
Percent refusals by neighborhood	29%	23%	17%	36%	39%
Percent refusals in reference cohort	$(25 + 8 + 3 + 10 + 7) / (86 + 35 + 18 + 28 + 18) = 29\%$				

Since contamination of the groundwater started, people have moved out of or into the study area, across the boundary of a household. Marriage, divorce, births, and deaths and other events changed the number and composition of Phase I households. Since no information was available on reference households *not* selected for Phase II, or households that refused to participate in Phase I, and since census data proved inadequate for estimating the size of the eligible study population, the denominator of the participation (or refusal) rate is unknown. It is thus difficult to estimate a participation rate on the basis of individuals, but a crude estimate can be extracted from Table 2.2, assuming that 1) the willingness of the respondent to participate in Phase II holds for the entire household, and 2) all members of the Phase I household are eligible for participation. Based on these assumptions, $14 \times 2.5 / (71 + 14) \times 2.4 = 13\%$ of the exposed Phase I population and $173 \times 3.1 / (281 + 173) \times 2.8 = 34\%$ of the reference Phase I population did not intend to participate in Phase II. Some people had changed their mind when Phase II actually started one year later. Two-thirds of the reference population eventually did *not* participate in the study, because the exposed/reference ratio of 2:1 was reached. The coverage of the population on the basis of *households* participating in Phase II is presented instead.

As shown in Tables 2.4 and 2.5, the refusal rate is 18% in the exposed area, and 29% in the reference area, close to the crude approximation of 13% and 34% from Table 2.2, and quite satisfactory. These rates are still an approximation, because they do not include tracked households in both the numerator and denominator of the rate.

It is difficult to define a tracked household, as is illustrated by the following example. One member of a current household left the area and was traced back as a tracked person. For

the contractor, this individual constituted a new household, but he was obviously still a member of his initial household. It becomes more complicated if two or more members of the household left and each established a household at a different address. The situation may be further confused if data on the tracked household are lacking because an interview was refused.

Thus, the numbers of households labelled "tracked" and "current" in both Table 2.3 and Appendix B do not correspond and should not be cross-referenced. For the contractor, each address successfully tracked was counted as a household, regardless of whether one person, the entire family, or even nobody at that address was interviewed, or whether the nucleus of the original household was still in the study area. The data provided in Appendix B, Table 3, suggest that only one exposed and six reference households refused to take part in Phase II. However, some or all of the people in those households might have been ineligible for participation in Phase II, as eligibility (date of moving in or out of the study area) could only be determined at the interview. This also holds for households which could not be located. In any event, Table 3 in Appendix B does not suggest that the refusal rate in the tracked population differs significantly from that in the current population. If this inference is correct, the current population refusal rate of 18% among the exposed cohort and less than 29% among the reference cohort may hold for the entire study population.

2.4.3 Data Gathering - Exposure Assessment

The water supply of study participants, including the water at the addresses to which people went after leaving the study area, was tested for VOCs. If people moved to an address with a municipal water connection, water quality data were requested from the municipal water authorities. All samples were tested for VOCs by the water laboratory of MDPH. The quality control procedures of this laboratory complied with those of the EPA.

Since no past contamination levels of residential wells were known, the assessment of individual exposure required an approach by approximation. Data from the monitoring program for the city wells of Battle Creek since September 1981 were used to estimate, through mathematical modeling, when the city wells became contaminated, and how VOC concentration levels have changed since that time. Results of this research were extrapolated to individual private wells in Verona Park to estimate, for *each person and each chemical*, when exposure began, the duration of exposure, and the total amount of VOC exposure during the stay in the study area (Chapter 4 and Appendix C). Further, an estimate was made of an individual's exposure to VOCs through showering and bathing in contaminated water.

When Phase II began, exposure of most people had ceased more than a year before. This eliminated the possibility of testing for VOCs or metabolites in serum or urine, as the biological half-life of VOCs at the prevailing contamination levels is a matter of hours. Furthermore, current or recent serum or urinary levels of VOCs or their metabolites are poor indicators of the exposure accumulated in the past (28, 84). As will be discussed in Chapter 4, the lack of monitoring well data for Springfield and Dowagiac has caused difficulties in solving exposure assessment problems.

Since most chronic diseases are multicausal, consideration has to be given to other chemicals which may induce the same kind of health events expected from VOCs. Therefore, the exposure assessment was expanded to include interview data on occupational exposure to VOCs and other chemicals, and blood tests on compounds persistent in the human body: chlorinated pesticides (DDT, chlordane and heptachlor), polychlorinated (PCB) and polybrominated biphenyls (PBB).

2.4.4 Data Gathering - Interview

All members of the selected or tracked households were interviewed in person, except for children, in which case one of the adults (preferably a parent) was asked to respond. As some of the questions might prove to touch upon sensitive matters, children between 10 and 18 had the option of responding in person. In the case of deceased persons, the next of kin were interviewed. Interviews were conducted at the residence by graduate students and recent graduates from the contractor's Department of Sociology. All interviewers were trained by the contractor for the specific purpose of this study. None of the 23 interviewers lived in the study area or knew the interviewees. Only three of the interviewers worked in both Phases I and II.

2.4.5 Data Gathering - Clinical Examination

Local clinics were set up for physical and clinical examinations of all study participants age 5 years and over. Clinic visits were arranged by appointment. Urine samples were processed at the clinic; blood samples were refrigerated to be tested in (sub)contract laboratories. The clinical examination consisted of measuring height, weight, upper arm circumference, temperature, blood pressure, and urinary beta-2-microglobulin (an indicator of toxic renal effects), and performing routine urinalysis and blood tests. The latter comprised routine hematology, including measures of glucose, hemoglobin A₁C, liver and renal function, acid phosphatase (for males age 50 and over), iron and iron binding capacity, cholesterol, and triglycerides. In addition, the MDPH tested serum for PCB, PBB, and chlorinated pesticides.

2.4.6 Data Gathering - Medical Records

In addition to collecting health data through interviewing, participants were asked, at each disease-specific question, for hospitals visited and physicians seen since 1970. This date was chosen: first, because it was very unlikely that the Verona area was contaminated by VOCs before 1970, since the generator of the pollution started its operations not earlier than 1968, and some time had to elapse for the spill to occur and for the spilled chemical to reach the wells; second, the further in the past the search for data, the more deficient people's recall capacity is and the more difficult it is to retrieve medical records. Hospitals and physicians named by the interviewees, and all hospitals in a wide area serving the study area were visited to retrieve and abstract medical records. This operation was carried out by qualified medical records personnel from the Borgess Medical Center (subcontractor). Copies of death records were obtained from the Center for Health Statistics, MDPH.

2.4.7 Data Gathering - Validation of Health and Laboratory Information

Care was taken to extend quality assurance and control procedures (QA/QC) to health data, in the expectation that the improved accuracy would minimize the occurrence of spurious exposure-event associations. The questionnaire, the review of medical records, and the clinical testing were designed to provide tools to evaluate and maximize the accuracy and credibility of the health data base (Chapter 5). Clinical chemistry testing was done under subcontract by the Borgess Medical Center, and was subjected to rigorous QA/QC guidance by the CDC (see Chapter 7) in addition to the subcontractor's in-house QA/QC procedures. Tests for PCB, PBB, and chlorinated pesticides, and tests for determining VOC levels in water samples were carried out by MDPH with QA/QC procedures that met EPA's requirements. Diagnosis validation criteria were established for each disease separately. An interviewee-reported disease needed support from medical record data, drug prescriptions, or clinical test results to be accepted as a case for analysis. This process is discussed further in Chapter 5.

2.5 Statistical Analysis

Epidemiologic analysis of a longitudinal study such as the Battle Creek Health Study is based on the comparison of an exposed cohort with a reference cohort, with regard to the frequency of health events occurring within the period of followup, and relative to the "person-time" incurred by the cohorts. The starting date of the followup period (TFU1) was January 1, 1970. For people who came to the study area or were born at a later date, TFU1 was the date of moving to the area or the date of birth. The closing date of the followup (TFU2) was the date of interview or the date of death. A valid case of a disease did not necessarily imply eligibility for inclusion in the statistical analysis. To be eligible, the date of diagnosis was required to be later than January 1, 1970 *and* later than the date a person moved to the study area. The comparison of the health experience in the exposed and the reference cohort can be expressed in a rate ratio or odds ratio (discussed in Chapter 6). A ratio above unity (1:1) suggests an excess of disease among the exposed; a ratio lower than unity suggests an excess of disease among the reference cohort.

The distribution of disease occurrence is rarely uniform, particularly in small populations, and a ratio of exactly 1.00 is very rare. Testing for statistical significance is one way, but certainly not a conclusive way, to judge to what extent chance might have caused the ratio to depart from unity. Such a departure may also result from insufficient control of confounding factors or from data inaccuracies, especially if inaccuracies are not randomly distributed. To decrease the possibility that incorrect conclusions might be drawn from spurious exposure-effect associations, the analysis was conducted in a number of ways, each reflecting a different point of view. Consistency in the results, across analytical models and exposure expressions, would support the validity of an observed association. Chapter 6 describes the analytical procedures used, and Chapter 9 provides the proper background for evaluating the results of the analysis.

Chapter 3 Toxicology

3.1 Introduction

Following risk assessment principles, the available toxicologic information on the VOCs was reviewed to identify health effects in humans possibly related to VOCs. Statistical associations of VOCs with health effects may indicate a causal relationship if these effects fit the toxicologic profile found in the literature. Otherwise, statistical associations are likely to be spurious or random. Toxic effects are not likely to be detected in a study of extremely low dose levels unless they have already been observed in studies of higher exposure levels.

Recent comprehensive literature reviews have been published by the EPA. In this chapter, the data on toxicity are based on these references (27 through 38), unless stated otherwise. An abundance of toxicologic information is available for the VOCs, except for CIS and 11DCA. For acute and subchronic toxicity data on 11DCA, the reference was a review document by the Netherlands Organization for Applied Research (92). The compiled information reveals that VOCs have very similar chemical structures and similar toxicologic profiles. CIS has been studied very little in animals or humans, but that which is known suggests close similarity to the other VOCs.

There is sufficient evidence that the intestinal absorption rate of VOCs is close or equal to 100%. Absorbed VOCs are metabolized in the liver at rates inversely related to the degree of chlorine substitution. Biotransformation of VOCs involves the microsomal cytochrome P450 system in the liver. The similar chemical structure of VOCs is reflected in the fact that the VOC metabolites are very similar to one another. Chloro-ethanols and acetic acids are metabolites of all VOCs, and epoxides and oxalic acid are metabolites of the ethylenes. Toxic effects are probably associated with adduct-formation of reactive metabolites with macromolecular components in the target organs. The VOCs' high affinity for lipids causes significant accumulation of VOCs in fatty tissues, but only as long as exposure is maintained. Since VOCs are rapidly cleared from the body after cessation of exposure, there is no value in testing for VOCs or metabolites more than a week after the exposure ends. Which toxic effect prevails is mainly determined by the magnitude of the dose, the route of exposure, and the species involved. The target organs and the associated effects are as summarized below:

3.2 Target Tissues and Systems

1. Central Nervous System (CNS): *Mild* symptoms and signs include headache, lightheadedness, loss of concentration, decreased mental performance, and fatigue. *Moderate* effects include exhilaration, loss of inhibition, mental fogginess or sluggishness, sleepiness, dizziness, nausea, and some loss of motor control and coordination. *More severe* effects include severe loss of balance and motor control, drowsiness, vomiting, inebriety, and eventually unconsciousness (narcosis), coma, and death.
2. Liver: Function disturbance and necrosis have been reported from accidents in anesthesia and high-dose poisoning, and have occasionally been fatal.
3. Kidney: Function disturbance and microscopic tubular necrosis are uncommon in man, but somewhat more frequent in animals at high doses.
4. Heart: As with many other solvents, VOCs appear to sensitize responses to epinephrine, a natural "stress" hormone, reflected in increased myocardial contractility, pulse rate, and

eventually ventricular fibrillation. The latter has caused most of the anesthesia deaths, and may explain most of the sudden deaths among addicts of VOC sniffing.

5. **Immune system:** A decreased humoral and cell mediated immune response has been observed in *in vitro* tests with blood of animals treated with TCE and 12DCA. The meaning of this is unclear, since the degree to which the results of *in vitro* tests bear on *in vivo* immune functions has yet to be established. No reports have been found of abnormal *in vitro* results for test systems of human origin, and no data from human or animal studies have suggested that VOCs may impair *in vivo* immune system responses. It should also be noted that positive *in vitro* tests on systems of human origin rapidly returned to normal after cessation of exposure.
6. **Local effects:** Upon local contact with undiluted or concentrated VOCs, tissue irritation or necrosis of superficial cell layers is likely to occur in the skin and mucous membranes of eyes, and respiratory and digestive tracts. Allergic reactions may develop, such as contact dermatitis and the "degreaser's flush". It should be recognized that contact dermatitis may occur with many chemicals, and even with solid metals such as nickel and chromium.
7. **Effect on the fetus:** There is no evidence of teratogenicity, although some VOCs have been shown to cross the placenta. Reproductive effects observed in animal tests may reflect an effect of exposure of the female to high levels of VOCs, or a direct effect on the fetus. These effects have included fetal loss, delayed ossification, and lower birth weight.

3.3 Mutagenicity and Cancer

Although some VOCs have been found mutagenic in a few bacterial tests, the effects on mammalian cells, the results of *in vivo* tests, and inconsistencies in the outcomes suggest that VOCs are, at most, weak or dubious mutagens.

One inhalation study on DCE involving rats, mice, and hamsters showed an excess risk of mammary tumors in female rats, and of renal cancer in male mice (59). However, repeat studies at the same institute with other strains of mice, and studies by others proved negative (59, 73, 82). PCE induced liver cancer in mice, but not in rats, in an oral study (67) which was criticized because overdosing resulted in high mortality in both mice and rats. The National Toxicology Program (NTP) repeated the study with four rat strains and one mouse strain, dosed by gavage and by inhalation. In its report on the inhalation part of the study (74), the NTP concluded that PCE caused an excess of mononuclear leukemia in the male and probably in the female F344/N rat, and liver cancer in mice. The meaning of the excess leukemia is not clear, since 1) doubling the dose did not increase the incidence, 2) control rats had a high incidence as well, 3) no leukemia was observed in the mice (spontaneously occurring leukemia was an uncommon disease in the mice), and 4) preliminary results of the gavage study did not show a leukemic effect in the rats (31).

The National Cancer Institute (NCI) conducted a gavage bioassay of TCA, and the compound was not found carcinogenic in rats and mice (64). The NCI considered the study inadequate, however, because of the very high, dose-related premature mortality. An inhalation study by Quast *et al* (80) was negative as well, but this result was challenged because too low a dose for too short a period was used (28). TCA was retested by the NTP, and preliminary results provided no evidence of an excess tumor incidence.

An oral study of 12DCA was positive for rats and mice, with tumors at multiple sites (65). However, there were flaws in the conduct of the study by the contractor (34, 60). Recent inhalation studies with much larger numbers of rats and mice per dose group did not reveal a

carcinogenic effect (60). A new study on three strains of rats and one strain of mice, exposed to 12DCA in water, is planned by the NTP (77). A TCE gavage study showed excess liver cancer in mice but not in rats (66); this study was disqualified by the NTP. A new NTP study was also found inadequate, primarily because of high mortality due to TCE toxicity (76). The review panel agreed that the value of conclusions drawn from an inadequate study is dubious, although the majority felt that TCE was toxic to the kidney, and that an observed excess of renal and testicular tumors in rats (not in mice) could not be evaluated due to the inadequacies of the study. CIS has not yet been tested in an animal cancer bioassay.

In summary, some VOCs appear carcinogenic in laboratory animals. Test results, however, are usually equivocal, not only because of deficiencies in the study design and performance, but also because an effect was usually observed in one species only. Positive responses were not confirmed by other studies in the same or different laboratories. There is no evidence of VOC carcinogenicity in humans.

3.4 Toxicity in Specific Exposure Scenarios

Human dose levels, and resultant toxic effects, are scenario-specific, as shown in the following examples.

1. Toxicity of deliberate exposure to VOCs:

VOCs have been used as anesthetics for many decades. For a full surgical narcosis, over 4000 ppm in air is usually required. Although the VOCs are safer than chloroform, deaths have occurred and, in the USA, these chemicals have been replaced with still safer compounds.

PCE in a single oral dose of up to 0.6 g/kg, has been used to cure intestinal infestation with worms. Reported transient side effects are faintness, giddiness, and drowsiness, but no liver or kidney effect.

Occupational exposure is mainly by inhalation. Occupational standards for air are on the order of several tens to hundreds of ppm, and are based on highest no-effect levels and on technological and economic feasibility.

TCE, TCA, and PCE have been abused for their CNS-effects, which include temporary euphoria. Sudden deaths have been reported, presumably resulting from an overdose causing ventricular fibrillation.

2. Toxicity of environmental exposure:

Prevailing VOC levels in ambient air and water are expressed in ppb, and the dose from food is on the order of micrograms per day. Environmental doses are thus about four orders of magnitude lower than those considered safe or acceptable in an occupational setting. Even when corrections are made for a longer exposure (24 hours a day rather than eight, seven days a week rather than five, and lifetime rather than a working life span of 45 years), assuming proportionality of discontinuous exposure, environmental levels are still about 1000 times lower than occupational levels of exposure. The general population includes children and elderly, who may be more sensitive to VOCs than workers. The EPA's Allowable Daily Intake has an included safety factor of 10 to protect sensitive populations. Thus, even at a continuous lifetime exposure with a safety factor of 10, using a conservative method to extrapolate discontinuous to continuous exposure, environmental doses are still one hundredth of what is considered to be safe occupationally. On the other

hand, studies reporting dose levels not causing adverse effects have involved relatively small numbers of workers, and may not have used the most sensitive methods of detecting toxic effects. The size of this study is larger than that of many occupational studies.

3. Toxicity of chloroform in drinking water:

Chlorination of drinking water causes the formation of trihalomethanes, one of which is chloroform. Its toxicologic profile is similar to that of the VOCs in this study (38, 80). Information on the chronic toxicity of other trihalomethanes is lacking, but the structural similarity suggests toxicological profiles similar to that of chloroform (75). Quantitatively, chloroform is approximately as toxic as TCA, TCE, and PCE with regard to the lethal dose and the highest chronic dose causing no toxic effect (compare Ref. 71 and below). Chloroform has been shown to be an animal carcinogen, although new findings suggest that the carcinogenic effect may have been attributable to the vehicle in which the agent was dissolved for intragastric administration (15). As explained in Chapter 4, inclusion of chloroform in the analysis is desirable from a methodological viewpoint.

4. Toxicity of exposure to multiple VOCs:

In evaluating the associations between exposure to VOCs and health events, a way has to be developed to convert the toxicity potential of multiple chemicals to that of a single compound. Since all VOCs are most likely acting through the same mechanism, and the target organs are the same, calculating the toxicity potential by addition is a reasonable approach. Addition is justified if the VOCs are equivalent in their toxicity potential per dose unit. Equivalency can be measured by comparing toxicologic parameters, such as: the LD₅₀ and TD₅₀ (the Lethal Dose or Toxic Dose in 50% of the animals); and the NOEL or the LOEL (No or Lowest Observed Effect Level).^{*} One human parameter, the occupational safety standard, is ill-suited for estimating a dose-equivalence, as it is usually compromised by non-toxicity issues.

To test the validity of the assumptions about the compound effect of multiple chemicals, the NTP initiated animal studies of the semichronic toxicity of chemical mixtures, including four of the VOCs in this study (77). Exposure is through drinking water at two dose levels. Until the results are known, inferences about the nature of interactions between VOCs (other than additive), and estimates of dose-equivalents remain speculative.

5. Effect of exposure to multiple chemicals other than VOCs:

With regard to concomitant exposure to unrelated chemicals, any chemical affecting the cytochrome P450 enzyme system might interact with VOCs. PCBs and phenobarbital are known inducers of these enzymes, and phenobarbital has been shown to increase VOC toxicity (62). Alcohol intolerance in males occupationally exposed to VOCs is known. The toxicologic profiles of PCB, PBB, and DDT differ from those of VOCs, and the LD₅₀ for these chemicals is much lower - that is, they are much more toxic than VOCs (14, 51). Rather than estimating a composite exposure with VOCs, these chemicals (PCB, PBB, and DDT) will be tested as separate variables for possible associations with health outcomes.

* Here and elsewhere in this report, NOEL and LOEL are used to indicate "adverse effect" rather than mere "effect". The proper terms to indicate adverse effects, NOAEL and LOAEL, are not in common use.

3.5 LD₅₀, LOEL, and NOEL

Of the parameters which may be used as the denominator for the dose-equivalents mentioned above, the LD₅₀ is the one most readily available. Dose-equivalents based on a LD₅₀ may also be associated with dose-equivalents for chronic diseases; this was demonstrated for cancer by Zeise *et al.*, who found a remarkably close correlation between the LD₅₀ and risk of cancer in animals (97). The LD₅₀ is determined by the toxicity and purity of the agent, animal species and strain, age, gender, the route and technique of administering the chemical, the duration of the exposure (in case of exposure by inhalation or dermal contact), and the duration of observation.

A search for data on the LD₅₀ of the seven VOCs and chloroform yielded widely varying values. Although the listing of LD₅₀ values in handbooks, government reports, and computerized databanks suggests mutual comparability, the original sources of the data revealed a striking lack of standardization, particularly with respect to the duration of the observation period, the dosing procedure, and the age and sex of the test animals. The effect of lack of standardization can be substantial, as reflected in a toxicity range as wide as 200 to 1800 mg/kg for DCE (36, 37). Only one laboratory (the Mellon Institute, University of Pittsburgh) standardized its testing protocol for the purpose of quantitatively comparing the toxicity level of different chemicals (86, 87). Unfortunately, only two of the VOCs in this study (TCE and 12DCA) and chloroform were included in this laboratory's program. No new reports on serial LD₅₀ testing have been published since 1969. The following is an overview of LD₅₀ data expressed in grams of VOC per kilogram of body weight for oral rat studies.

TCA	12.3 and 14.3 for male rats; 10.3 and 11 for female rats
TCE	7.2 for male rats (Mellon Institute)
PCE	13 and 8.85
11DCA	1.12 and 0.725
12DCA	0.68 and 0.77 (the latter for male rats; Mellon Institute)
CIS	0.7 for a cis/trans mixture of unknown composition
DCE	various values ranging from 0.2 to 1.8
Chloroform	3.2 for male rats (Mellon Institute)

Considering that the above values suffer from lack of standardization, a rough approximation yields rat LD₅₀ dose-equivalent values of: 10 for the higher chlorinated compounds (TCA, TCE, and PCE); 1 for the other VOCs; and an intermediate value for chloroform. The 10:1 ratio was also found for the LD₅₀ of TCE and 12DCA tested by the Mellon Institute in a standardized fashion (89). The LD₅₀ levels were sufficiently high to exceed the saturation limit (metabolic saturation was observed for at least TCA, TCE, and PCE). The 10:1 ratio is probably irrelevant to the present study, since the low Battle Creek exposure levels preclude acute effects. The very much lower doses of NOEL studies, and the still lower VOC levels in drinking water, are well below the saturation limit for humans.

As measures, the TD₅₀, the LOEL and the NOEL are theoretically to be preferred over the LD₅₀, since they focus on a specific health event rather than death. They are lower on the dose scale and thus closer to the dose levels of concern in this study. The search for TD₅₀ data on effects other than cancer yielded no useful information. There was not a single effect for which

a TD₅₀ was found for more than two VOCs. With regard to cancer, it was noted above that the evidence for such an effect is equivocal, and involved four chemicals only. The data on the LOEL was equally disappointing, as most studies were not designed specifically for this purpose. The reported LOELs pertained to different health events, and lower dose levels were usually not studied. The NOEL could have been more useful, since human data are available. However, the lack of standardization was very apparent, especially with regard to the sensitivity of the toxicity indicators. Since no oral NOELs for humans have been reported, it is necessary to convert an inhalation NOEL (usually expressed in ppm of chemical in air) into a NOEL of absorbed mg/kg/day.

$$\text{VOC absorbed (lung)} = (\text{NOEL} \times \text{CF} \times \text{Hr} \times \text{D} \times \text{R}) / (7 \times 70) \text{ mg/kg/d} \quad (\text{Eq 1})$$

- CF = molecular weight/24.45, which converts ppm into mg/m³
 (3.96 for DCE and CIS, 4.05 for 11DCA and 12 DCA,
 5.37 for TCE, 5.45 for TCA, and 6.78 for PCE)
- Hr = hours exposure per day
- D = days exposure per week
- R = proportion of inhaled chemical retained in the body
- 70 = average weight of a human adult

Equation 1 is based on an average respiration volume of one cubic meter per hour, and on the implied assumption that discontinuous exposure can be converted to continuous exposure by direct proportionality. This is a common assumption in regulatory risk assessment, although it is probably invalid because of recovery processes in the dose-free periods. A review of pertinent literature revealed that the pulmonary retention of VOCs in men is approximately 55% (range 36-75%) for CE, 70% (range 60-80%) for PCE, 49% (range 38-60%) for DCE, and 28% (range 25-30%) for TCA (20, 28, 30, 33, 36). It is commonly assumed that the percent absorption is constant (28, 30, 33), although increasing the level of exposure to DCE from 25 to 300 ppm has been shown to reduce the uptake after three hours exposure from 77% to 50% (20). Thus, at the much lower ppb level, the percent absorption in the lungs may be higher than expected from the above values.

To calculate an oral NOEL for CIS, 11DCA, and 12DCA, for which no absorption rates have been published, a rate of 50% was assumed (a rounded average of the rates known for the other VOCs). In the case of animal studies, no corrections were made for interspecies differences in metabolism, respiration rate, lung surface/body weight ratio, or sensitivity of health effect indicators, due to lack of information. No NOEL- or LOEL-yielding chronic or subchronic studies on CIS were found. In selecting NOELs from the reviewed studies, preference was given to studies that included higher doses demonstrating some toxic effect. Studies not including such doses are prone to underestimate the true NOEL. An overview of the calculated NOELs follows:

TCA:

Human: Two occupational inhalation studies yielded a NOEL. The highest NOEL was 250 ppm (8 hr/weekday), equal to an oral NOEL of 31 mg/kg/d. The effect parameters included an electrocardiogram, clinical chemistry, blood pressure measurement, and a health questionnaire emphasizing cardiovascular and CNS dysfunction (53).

TCE:

Human: Several occupational inhalation studies yielded a NOEL. The highest NOEL was 25 ppm (8 hr/weekday), equal to an oral NOEL of 6 mg/kg/d. The effect parameters included a health questionnaire and clinical examination of CNS symptoms and behavioral performance, urinalysis, hematology, and limited blood chemistry (91).

Rat: An inhalation NOEL was found of 55 ppm (8 hr per weekday), equal to a human oral NOEL of 13 mg/kg/d. The effect parameters included clinical chemistry and hematology (52).

PCE:

Human: In several occupational inhalation studies the highest tested exposure yielding a NOEL varied from 18 ppm to 32 ppm (8 hr per weekday). The 18 ppm level equals an oral NOEL of 7 mg/kg/d. No higher exposure was tested. The effect parameters were neurological tests, blood glucose, hematocrit, and behavioral performance (94).

Rat: An inhalation NOEL was found of 70 ppm (8 hr per weekday), equal to a human oral NOEL of 27 mg/kg/d. The test parameters were histology and blood tests including bilirubin, glucose, calcium, and a WBC count (17).

11DCA:

Human: No studies yielding a NOEL.

Rat, Guinea Pig & Rabbit: An inhalation NOEL of 1000 ppm (6 hr per weekday) was found, equal to a human oral NOEL of 124 mg/kg/d. The NOEL in cats was 500 ppm, equal to a human NOEL of 62 mg/kg/d. The parameters were hematology, clinical chemistry, and histology (46).

12DCA:

Human: No studies yielding a NOEL.

Rat, Cat, Rabbit, Guinea Pig: An inhalation NOEL of 100 ppm (6 hr/weekday) was found, equal to a human oral NOEL of 12 mg/kg/d. The parameters were hematology, clinical chemistry, and histology (46).

DCE:

Human: No studies yielding a NOEL.

Rat & Dog: An oral NOEL was found of 10 mg/kg/d (no conversion to human oral NOEL is needed, since these are oral studies). The effect parameters included histology and clinical chemistry (82).

Chloroform (CHL):

Rat: An oral 3-months study revealed a NOEL of 30 mg/kg. Parameters were liver and kidney toxicity (80).

In its guidelines, the EPA has followed the concept of additivity of dose-equivalents in assessing the toxicity of mixtures (26). The EPA's concept of an equivalent was the dose divided by a measure of "acceptable level", usually the Adjusted Acceptable Daily Intake (AADI). The AADI is essentially the NOEL at a water intake of two liters a day, divided by a "safety factor" of 10, 100, or 1000 (depending on the quality and kind of data) and a factor to adjust for the extent of exposure to the chemical from other sources. A weakness specific to the use of the AADI rather than the NOEL is that the selection and application of safety factors have now obtained a regulatory policy character. Although using a safety factor can be defended scientifically, determining its magnitude is, in practice, almost entirely subjective, with little or no reference to biological facts and the magnitude of the error caused by deficient data. The same data set may, therefore, give rise to widely different AADIs, dependent on subjective and inconsistent evaluation of study results yielding a NOEL. This renders the AADI a parameter even less suitable for comparative purposes than the NOEL.

The 10:1 ratio observed in the LD₅₀ for higher and lower chlorinated compounds has not been confirmed for the NOEL. The differences among the NOELs for the chemicals seem large; however, the studies yielding the NOELs were not standardized for test conditions and parameters of effect. Also, it was necessary to convert inhalation data into an oral NOEL, using (in some instances) assumed rather than observed pulmonary retention rates. It was assumed that equal amounts of VOCs absorbed in the lungs and in the gut are equitoxic; this may be an incorrect assumption because of different pathways with respect to passage through the liver. Finally, the manner in which animal NOELs relate to human NOELs is unknown. Taking all of these uncertainties into account, the above data suggest that there is little quantitative difference among the chronic toxicities of the VOCs (plus chloroform), although the high NOEL for 11DCA may be accurate, since it concurs with the reported very low toxicity of this compound. Considering the irrelevance of LD₅₀ data for low-dose chronic effects and the approximately 1:1 dose-equivalence of the NOELs, it appears appropriate to stipulate that the composite exposure to multiple VOCs and chloroform is equal to the simple sum of the chemical-specific exposures.

Chapter 4 Exposure Assessment

4.1 Introduction

In a cohort study such as this, the first requirement is to define exposure and ascertain qualitatively and quantitatively the individual exposure status of cohort members. Typically, cohort studies are conducted in an occupational or clinical setting with *a priori* knowledge of the exposure status of the study participants. This permits assignment of individuals to exposed and reference cohorts. This study is quite different, since the only *a priori* knowledge concerned a contaminated aquifer. Cohorts were formed *expecting*, not knowing, that the exposure status of an area cohort would coincide with individual exposure status. At the time the study commenced, it was not known when exposure had begun -- another distinction between this study and classic cohort studies. Chronic disease studies typically focus on studying the relative impact of risk factors; an environmental study, however, estimates the effect of known causes. Thus, it is crucial to estimate the exposure level (C) at various points in time (T), as well as the date that individual exposure began. This enables the validation of a correct exposure-disease sequence and an evaluation of the dose-response relationship. Failure to estimate the dynamic exposure characteristics would considerably increase the likelihood of spurious statistical exposure-effect associations.

4.2 Estimating Individual Exposure to VOCs from Drinking Water

Conducting a study in a situation characterized by lack of individual exposure data requires the development of new methods for assessing individual exposure. The method developed for this study involved the generation of a mathematical model for individual exposure. Although lack of calibration is inherent in novel methods, modeling is preferred to its alternatives: 1) the conventional but erroneous assumption that exposure was constant during the period of residency in the exposed area; 2) the equally erroneous assumption that the exposure level found in a single and current sample can be extended to the entire cohort; or 3) that living in a contaminated area is synonymous to being exposed.

The method for estimating the date a well became contaminated (T_1), how a chemical's concentration changed over time, $C = f(T)$, and the total accumulated exposure (TAE), is described in Appendix 3. In essence, the method utilizes data from the available time series of polluted city wells to estimate $C = f(T)$, accounting for differences in well properties. Once estimated for each chemical separately, $C = f(T)$ is applied to the single-sample data from individual residential wells, to estimate T_1 for individual wells, and a TAE for individuals. TAE is derived from $C = f(T)$ by integrating the area under the concentration-time curve (CT-curve) between the dates that exposure started and ceased. The following equations (equations 4 and 6 in Appendix C) for T_1 and TAE were developed. In these equations, C_s is the sample concentration and T_s the date a private well was sampled; b is the slope or the chemical-specific coefficient of time T in the $\log C = T$ regression model underlying equations 2 and 3. Time values are in months since 1-1-1970.

$$T_1 = T_s - (\log C_s/b) \text{ months since 1-1-1970} \quad (\text{Eq 2})$$

$$\text{TAE} = (\exp(b(T_s - T_1)) - 1)/b \text{ ppb-months} \quad (\text{Eq 3})$$

Since the date of moving to the study area (TIN) was often later than T_1 , and TSTOP (date exposure stopped) was usually later than T_s , equation 3 can be generalized as follows:

$$\text{TAE} = (\exp(b(\text{TSTOP} - T_1)) - \exp(b(\text{TIN} - T_1)))/b \quad \text{ppb-months} \quad (\text{Eq 4})$$

The above equations are based on values for b estimated by modeling Battle Creek city wells. With regard to the applicability of b to the various study areas with contaminated wells, there are several local scenarios:

- 1) Private wells in Verona Park *west* of the source of pollution (Area 2, Figure 1.2) are in the main pathway of the groundwater flow from the source to the river and the city wells. The model $C = f(T)$ was developed specifically for these wells, and calibration of the model with data from private wells sampled twice gave quite satisfactory results. The T_1 estimates were compatible with crude estimates based on groundwater velocity.
- 2) Private wells in Verona Park directly *north* of the source (Area 3, Figure 1.2) are in the outer zone of the plume of contaminants, while the area *south* of the source (Area 4, Figure 1.2) is upstream. Although the increase in VOC-levels in Area 3 would still depend on the well characteristics used to develop $C = f(T)$, the slope b is probably smaller (slower increase over time) than that of the city wells. In Area 4, progression of the plume is by diffusion only, and using $C = f(T)$ for these wells would require expansion of the model with components reflecting the negative effect of an upstream location. Since none of the city wells are upstream, there are no data allowing the development of such components. In Areas 3 and 4, b is likely to be smaller than b for the city wells, rendering TAE estimates probably too low, and T_1 estimates too early. VOC levels in these areas were very low, and were measured more recently compared with Area 2. Most values were at or slightly above 1 ppb, and the sampling date for most VOCs was equal to T_1 . Under these conditions, errors caused by applying equations 2 and 4 are negligibly small, and even if b were much smaller, the TAE for Areas 3 and 4 would still remain at the lower end of the range of values measured for Area 2.
- 3) No point source was identified in Springfield. Accordingly, there is no rationale for applying the $C = f(T)$ model to wells in this area. The VOC concentrations in these private wells were low, and there is no major VOC handling company nearby. The VOCs probably were, and still are, spilled into the aquifer at multiple sites by domestic waste-handling activities, small leaking tanks, or small-scale industrial waste operations.
- 4) In Dowagiac, a point source was identified, but there were no historic data enabling the development of a $C = f(T)$ model specific to that area. Using the b -values from the Battle Creek wells may result in large errors in T_1 and TAE, if the real b -value significantly differs from the assumed value.

In the epidemiologic analysis, it was assumed that in Springfield and Dowagiac the concentration at time T_j was constant since 1-1-1970, the date of the start of the followup period. TAE for residents in these areas is then the simple product of C_j and the time elapsed since 1-1-1970 (or TIN, if that date was later) until TSTOP. For Springfield residents, the assumption that C was constant over T since 1-1-1970 might result in a substantial overestimation of TAE. Any other assumption, however, would be at least as potentially erroneous. As long as no point source can be identified, there is a real probability that the pollution resulted from a long-standing practice of small-scale waste disposal with spillage into the aquifer, eventually leading to a steady-state low contamination level. The impact on the analysis of assuming constant levels since 1970 is very limited, since the cohort size is just 16 people, only three of whom accumulated a substantial TAE.

For Dowagiac residents, errors could be large due to the high VOC-levels observed in that area. The assumption that the contamination level has been constant since 1970 has some support

in that a point source has been identified; it had been in operation for decades with inadequate solvent storage facilities. There was ample time for the VOCs in the aquifer to reach an equilibrium. The distance from the source to the wells is similar to that in Verona Park. As the concentration levels in Dowagiac are approximately one order of magnitude higher than those in Verona Park, it is within expectation that T_1 for Dowagiac wells is many years before 1977, the earliest estimate for Verona Park wells (see Table 4.1). The impact these assumptions may have on the evaluation of exposure-effect associations will be tested. This can be done by an additional evaluation of exposure-effect associations in a restricted population comprising the Verona exposed and the Calhoun county reference cohorts only.

Table 4.1 Summary of b -values (regression slope) and resultant estimated T_1 values for Verona residential wells (Equation 2). T_1 is expressed in terms of month and year.

	TCA	11DCA	12DCA	DCE	CIS	TCE	PCE	Total VOC
b -value	0.1536	0.1034	0.1048	0.0795	0.1667	0.1434	0.1493	
earliest T_1	05-79	10-77	03-77	03-77	08-77	04-78	01-79	03-77
latest T_1	11-82	07-83	08-82	05-82	11-82	09-82	07-83	07-83
median T_1	05-81	01-80	11-79	04-80	04-80	01-80	06-80	

Table 4.1 depicts the range of T_1 estimates for Verona Park residential wells. Fifty percent of the residential Verona wells (median T_1) became contaminated not earlier than November 1979. The range of T_1 is 6 years (1977-1983), with the earliest T_1 estimates within the main pathway of the groundwater flow, and the latest estimates for wells outside the main pathway. Neighboring wells showed a very much narrower range (Figure C.3, Appendix C). The range of earliest T_1 for different VOCs is approximately 2 years, indicating that spilling did not start at the same time for all chemicals. As the various chemicals were handled and stored at different times and in different amounts, depending on supply and demand, the time and volume of spillage (thus T_1 and b) would be expected to vary substantially. This is consistent with the finding of differences among observed VOC-specific concentrations in the same sample from city wells, and among VOC-specific T_1 values in the few cases in which a city well had an observed rather than estimated T_1 .

Key elements in the assessment of individual exposure are the dates of: moving into (TIN) or out of (TOUT) the study area; discontinuing the use of VOC-contaminated water (TSTOP); and first diagnosis (TDIAG) of a disease. The relation between these points in time and TAE (the area under the CT-curve) is illustrated in Figure 4.1 below, for three different scenarios resulting in three different exposure estimates E1, E2, and E3. Estimating TAE based on so many points in time, different for each chemical and disease, is a very complex procedure for which a computer program was developed (Appendix I) to prevent errors or biases based on foreknowledge of individual exposure and health status.

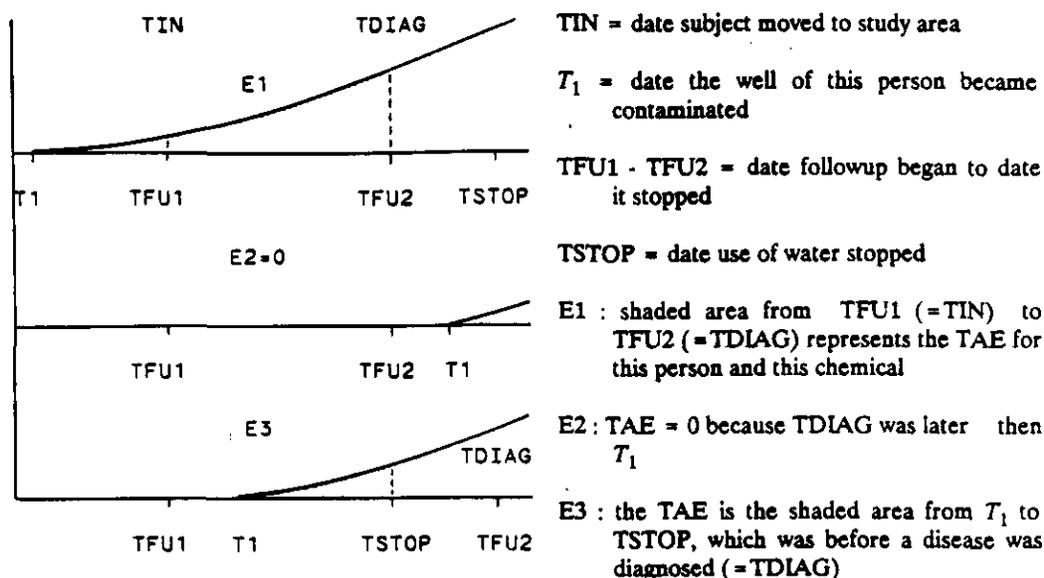


Figure 4.1 Determinants of scenario- and chemical-specific TAE. E1, E2, and E3 are the shaded areas under the curve, and represent TAE estimates for 3 individuals, each with a set of time events. The Y-axis of the curve represents the concentration level, and the X-axis the time of followup.

When dealing with exposure to multiple chemicals with the same toxicological profile and route of entry, the most logical and effective approach to the analysis of exposure-effect associations is to use dose-equivalents to convert the multiple exposures to a single one. The evaluation of NOELs following chronic exposure (Chapter 3) suggested an approximate dose-equivalency per weight unit absorbed, rendering the accumulated exposure to all VOCs combined (TAEVOC) equal to the sum of VOC-specific TAEs (equation 5).

$$\text{TAEVOC} = \text{TAETCA} + \text{TAEPCE} + \text{TAETCE} + \text{TAEDCE} + \text{TAE11DCA} + \text{TAE12DCA} + \text{TAE CIS} \quad (\text{Eq 5})$$

As described in Chapter 3, chloroform was found to be equitoxic at low doses, and TAEVOC can be adjusted for chloroform by simple addition of its TAE (TAECHL), resulting in an alternative exposure value TAEVCL = TAEVOC + TAECHL.

As mentioned earlier, the pattern of concentrations of VOCs in well water (Table 1.1), gives a misleading impression of individual exposure levels. Table 4.2 shows the effect of applying to these concentrations the information on the actual period of residence, TSTOP, $C = f(T)$, and T_1 . The estimates in this table result from equations 4 and 5, and are based on the actual data on TIN and TSTOP compiled in Phase II. They are computed as if no disease occurred between T_1 and TSTOP.

The highest TAEVOC levels found were in Verona Park West and in Dowagiac. As can be deduced from Table 1.1, these high values were caused mainly by TCA and TCE in Dowagiac, and CIS in Verona West. Even if the date at which exposure in Dowagiac reached a state of

equilibrium were advanced to 1975, participants in this study area would still have the second highest exposure levels. The negligibly low exposure level in Verona areas North and South is better illustrated in Table 4.2 than in Table 1.1.

Table 4.2 Summary of individual Phase II estimates of exposure to all 7 VOCs combined (TAEVOC) in ppb-months, using equation 5 for Verona, and $C_1 \times$ duration for Dowagiac and Springfield; calculated as if no disease occurred between T_1 and TSTOP.

	West	Verona Park North	South	Dowagiac	Springfield
Highest TAEVOC	90x10 ⁶	42	368	1.9x10 ⁶	16065
Median TAEVOC	3680	42	26	45123	9128
Number of people TAEVOC>0	134	6	17	50	16
Number of people TAEVOC=0	10	0	14	4	0

Table 4.3 Distribution of exposure values, expressed as the sum of chemical-specific concentrations in a current water sample and as TAEVOC accumulated during the period of residence in the contaminated area.

Current Water Sample			Concentration-Time Integration		
People	%	Sum VOCs (ppb)	People	%	TAEVOC (ppbmonths)
0	0	0	28	11	0
138	55	1 - 9	49	20	1 - 99
60	24	100 - 1,000	33	13	100 - 1,000
43	17	1,000 - 10,000	58	23	1,000 - 10,000
10	4	10,000 - 100,000	47	19	10,000 - 100,000
-	-	-	25	10	100,000 - 1,000,000
-	-	-	11	4	1,000,000 - 10,000,000

Table 4.3 clearly illustrates the effect of the individual exposure assessment using all available data, rather than assuming that a simple concentration in a current sample is an adequate measure of exposure. In fact, 11% of the people in the contaminated area were never exposed because they came into the area after their well was disconnected, or because they moved out before contamination of their well started. Further, the period of residence relative to T_1 is most influential in estimating the TAEVOC incurred during residence. This is reflected in the fact that exposure calculated on the basis of current concentrations alone, would assign 55% of the people to the lowest exposure category, compared to 20% using the TAEVOC value.

4.3 Conversion of Exposure to Dose

TAEVOC and TAEVCL are no more than a measure of "available" exposure. An individual with a positive TAE of any amount may still have been unexposed if no fresh tap water was consumed at home. To calculate a DOSE, the volume of water intake needs to be known. Interview responses showed a mean of 5.6 glasses of unheated tap water consumed per day at home (as plain water or in cold drinks) for the total population (median for exposed cohort 4, median for reference cohort 5). The highest number (64 glasses a day) was reported in the Verona exposed cohort, the second highest (48 glasses) in the Barron Lake reference cohort. Such values, not counting water intake on the job or in hot drinks and food, do not appear credible, but they were accepted for analysis in the absence of information which could be used for validation. Since there was no significant difference in the (abnormally) high values between exposed and unexposed cohorts, data inaccuracies are not likely to affect the analysis of the effect of VOC exposure. If WATER is the number of 8 ounce glasses of water drunk a day, $DOSEVOC(VCL) = TAEVOC(VCL) \times WATER$. The accumulated DOSEVOC(VCL) can be converted into mg total VOC/day with the following equation:

$$DOSEVOC(VCL) = \frac{TAEVOC(VCL) \times WATER \times 0.00024}{(\text{duration of exposure in months})} \text{ mg total VOC/day} \quad (\text{Eq 6})$$

4.4 VOCs in Drinking Water of Comparison Areas

When selecting comparison neighborhoods, it was assumed that the wells were not contaminated. In Phase II, however, all wells of the reference cohort were tested for the presence of VOCs, and none contained VOCs.

4.5 VOCs in Current Water Supplies of Former Residents of the Study Area

In agreement with risk assessment principles, the study protocol also called for testing of the current water supply of persons who had left the study area, in order to determine their exposure status at the new address. If people had moved to a new dwelling with a well, a sample was taken and tested for VOCs by MDPH. If people had moved to a city with a community water system, the MDPH requested the city and water authorities to provide copies of the results of recent water testing. None of the samples of private wells outside the study area contained VOCs. Only one family, former residents of a comparison neighborhood, moved to a city with a VOC-contaminated water supply. The followup of these residents stopped at the date they moved, that is, TFU2 = TOUT. The information received from city water authorities is summarized as follows:

Six cities provided full data on all seven VOCs: no VOCs were detected at a range of detection levels from 0.1 - 2 ppb.

One city stated that no VOCs or other organic compounds were found in eight water samples taken in 1986, but no individual sample data were provided.

One city provided the results of a 1978 EPA survey (detection limit 0.5 ppb). No VOCs were detected, but there was no testing for 11DCE and 11DCA.

One city provided test results showing no detectable levels of organic compounds. VOCs were among the chemicals tested, but it was not clear from the format of the results whether the tests included all listed chemicals.

One city did not test for organic compounds, as the community was too small with no industrial activity upstream. Its water source was strictly melted snow.

One city reported substantial contamination with TCE (maximum of 95 ppb), TCA and PCE were found at trace level (0.1 and 3 ppb), while 11DCA was found at a level of 18 ppb.

One city did not provide the requested information.

4.6 VOC Exposure Through Bathing and Showering

Conceptually, a combined exposure for multiple routes of entry can be achieved by adding TAEs from washing, bathing, and drinking. This would require estimating route-specific dose-equivalents, however, which is complicated by problems for which there are yet no adequate solutions. Brown *et al.* (13) claimed that skin absorption of VOCs from water is directly proportional to the chemical concentration. In other words, the ratio of absorbed amount to concentration is a constant (permeation constant). This claim appears incorrect. It ignores the effect of water/skin lipids and lipids/blood partition coefficients which differ between chemicals. Actually, Brown's claim results from improperly rounding the ratios. Adding just one decimal place reveals a consistent decrease of all "constants" with decreasing concentration; that is, the constant for ethylbenzene drops as much as 28% if the chemical concentration decreases by 29%. As VOC-levels in the Battle Creek Study are four orders of magnitude lower than those in the studies reviewed by Brown, this dependence of the permeation rate on the concentration might result in no observable absorption at these low levels. This lack of skin absorption data at low concentrations prohibits the estimation of dose-equivalents for this route.

The estimation of dose-equivalents for exposure through inhalation gives rise to other problems. Human pulmonary retention rates were found for only four of the seven VOCs (Chapter 3). Some have claimed that these rates are independent of the VOC-concentration in the air, but Dallas *et al.* (20) showed an uptake inversely related to the exposure level. As the experimental concentrations were several orders of magnitude higher than those prevailing in a bathroom, and the bathroom exposure is too short to reach an equilibrium compatible with the water/air and the air/blood partition coefficients, the percent uptake in the lungs in a residential setting is likely to be quite different from that seen in work place or experimental scenarios.

Air concentrations resulting from evaporation depend on: water and air temperatures; the duration and frequency of bathing or showering; the degree of bathroom ventilation; and the volume of the bath water. None of these data were available or could have been made available as part of this study. They can be expected to differ not only between people, but also for the same individual from day to day. McKone (61) reported preliminary results from a new model for predicting the absorption of VOCs in bathrooms. The model suggests that the absorbed amount could be as much as six times the amount absorbed when drinking two liters of contaminated water a day. This model was not calibrated with real data. It used many untested assumptions in extrapolation from the high theoretical concentrations down to the much lower actual concentrations prevailing in bathrooms.

In summary, there is evidence that with decreasing VOC concentrations, the absorption in the lungs increases, whereas the skin absorption decreases. This renders extrapolation models assuming constancy of percent uptake invalid. Therefore, due to the lack of essential physiological information which could be used to estimate dose-equivalents, exposure to VOCs through skin absorption and inhalation could not be incorporated into TAE or DOSE estimates. To ameliorate this problem, the secondary exposure was transformed into a new variable, WASH, equal to:

$$\begin{aligned} \text{WASH} &= \text{BATH} + \text{SHOWER} && \text{(Eq 7)} \\ \text{BATH} &= (\text{baths/week}) \text{ times minutes per bath;} \\ \text{SHOWER} &= (\text{showers/week}) \text{ times minutes per shower.} \end{aligned}$$

The variable WASH should be interpreted as a semiquantitative or an ordinal variable. On the one hand, it preserves individual variability; on the other hand, it ignores physico-chemical differences between the skin and lung routes. If a sufficiently large number of cases is available, BATH and SHOWER may be used as two distinct variables in a multivariable model, thus accounting for possible differences in dose-equivalents between the routes of exposure.

4.7 VOC Exposure from Secondary Sources

Since there were no provisions for estimating VOC levels in the work place, data on occupational exposure were limited to responses to an interview question and to a list of chemicals. Respondents were provided with categories of chemicals that could comprise VOCs. The response was sufficiently detailed to create new variables with some bearing on VOCs, as described below:

- WORK1 = Solvents and degreasers; these products typically consist of VOCs, and interviewees often specifically mentioned some of the VOCs.
- WORK2 = Other liquid products, comprising petroleum products, highly volatile compounds such as acetone, paints and paint removers, cleansing agents, printers ink, etc. Some of these products may contain VOCs (e.g., leaded gasoline contains 12DCA);
- WORK3 = Pesticides; these may be dissolved in VOC-containing compounds.

Variations such as pooling WORK1 and WORK2, were included in the analysis.

Questions on possible VOC exposure through consumer products yielded a wide variety of products without evidence of cohort-preference. Frequently, not all of the members of the same household had the same responses. Using the responses as evidence of exposure might, therefore, lead to misclassification. Finally, a search for qualitative and quantitative data on the contents of the many consumer products, most of which were not reported by brand name and none by concentration, would have required resources far exceeding those available for this study. The exposure to VOCs through consumer products was hence disregarded in the analysis.

4.8 Exposure to Other Chemicals

As discussed in Chapter 3, chloroform has a toxicologic profile similar to that of the VOCs. Since exposure to chloroform involved primarily the Battle Creek City reference cohort, inclusion of chloroform may confound exposure-disease associations. Chloroform levels in Battle Creek City tap water have ranged from 2 to 15 ppb (since 1982) with an average of 6 ppb. In view of the stable levels over a monitoring period of 5 years, it is reasonable to assume that the levels have remained stable since chlorination started before World War II. Based on equitoxicity considerations, it was concluded that TAEVCL (TAEVOC + TAECHL, Section 4.2) may be a valuable alternative measure of exposure. Because of the stable chloroform level in tap water, TAECHL is simply the concentration (6 ppb) times the duration of exposure to chlorinated city water. Since Verona Park, Dowagiac, and Springfield residents became connected to city water mains or had used bottled water after the groundwater contamination was detected (the bottled

water was tested, it contained as much or more chloroform as city water), TAECHL also was computed for these residents, although its magnitude was negligibly small compared to TAEVOC.

Blood samples were tested for PCB, PBB, and metabolites of chlordane, heptachlor, and DDT. These compounds and metabolites are persistent in the environment and body tissues, and share most of the toxic effects of VOC-exposure (14,51). The questionnaire was another source of information on household or occupational exposure to pesticides. PBB in serum is common in Michigan residents since the accidental introduction of this fire retardant into the food chain in 1973 (49). The potential role of these chemicals in association with health effects will be considered, since exposure may differ among the cohorts.

Chapter 5 Health Assessment

5.1 Introduction

To minimize the potential for an inaccurate health assessment which could cause spurious associations between disease and exposure, it was decided from the outset that health data collected should be of a quality consistent with risk assessment principles. Information on an individual's past and current health status was obtained by means of interviewing, retrieving and abstracting medical records, and clinical examination. The following sections describe the procedures used in assessing the individual health status in Phase II of the study.

5.2 The Questionnaire

To minimize the potential for misunderstanding at the interview, the draft questionnaire and consent form for release of all medical records were presented to the contractor, the Kercher Center for Social Research, for pilot testing and rephrasing. The interviews were conducted at home by appointment. Health questions of key interest were repeated in different formats and wordings, as shown below:

General question:	What are the <i>current</i> (at interview) diseases ?
	What is the <i>current</i> (at interview) medication ?
Specific question:	As of 1970, did you have (name of disease) ?
Specific question:	Did you have (disease) before 1970 ? When ?
Last organ-specific question:	Other diseases not mentioned earlier ?
In case of diabetes:	Did you use insulin ?

Additional questions were asked at the clinic:

- a) What are the current diseases not mentioned earlier ?
- b) What medication was taken, not mentioned at interview ?
- c) When was the medication for hypertension last taken ?

Differences in responses to such questions could point to recall deficiencies. Each question addressing the occurrence of a disease also included a request for the names and addresses of the hospital and physician involved, as well as the date of first diagnosis. To reduce misunderstanding and maximize recall, diseases were grouped in broad categories. Thus a category which started with "What about blood and vascular problems", included questions about anemia, leukemia, hypertension, and other diseases. As an aid in obtaining information on past and current medication, the interviewee was shown a list of various categories of drugs, e.g. pain killers, drugs for treatment of increased blood pressure, antibiotics, etc. People were also asked to show the interviewer copies of drug prescriptions, if available, and the labels or containers of prescription drugs, to improve the accuracy of the information on drug names and dates of prescription.

At the completion of an interviewing day, the interviewers reported directly to the contractor's Manager of Field Operations to submit the questionnaires, and to discuss problems that might have occurred. Respondents were contacted again if it was deemed necessary to clarify a problem. To reduce errors in coding, the interviewer's task did not include coding. Coding was done by other personnel on preformatted and computer-read forms, which eliminated key-punch errors. The health-related part of the interview yielded 357 computerized data variables. The responses to 50 questions on names of prescription or over-the-counter drugs,

"other" diseases and symptoms, and names and addresses of physicians and hospitals, were not computerized. At the CEHIC, these responses were screened for mention of diseases and medications of interest to the analysis. Responses were validated in the context of the overall picture arising from all medical records and clinical test results.

5.3 Medical Records

All medical records obtained from hospitals and private physicians (Chapter 2) were scrutinized for diagnoses. The diagnosis, ICD code, and date of discharge or visit were recorded on medical abstract sheets. In the case of private office records, the date of the visit to the office was sometimes missing, and the diagnosis was often merely a repeat of the patient's complaints. The hospital record data base comprised diagnoses as worded in the medical record, which might differ from the actual disease. For instance, a patient with coronary heart disease since 1970 might have multiple entries in the data base between 1970 and 1985, mentioning chest pain, angina pectoris, myocardial ischemia, arteriosclerotic heart disease, ischemic heart disease, abnormal ECG, etc.

All hospitals in a wide area serving the study population (see Appendix B) and, in addition, the hospitals and physicians specifically named by participants, were visited to retrieve and abstract medical records. This was done under subcontract by certified hospital registrars from the Borgess Medical Center, supervised by that Center's Project Manager. Using local registrars was considered less desirable, since it could affect the uniformity and quality of record searches, abstracts, and coding of diseases. Hospitals and physicians outside the region were asked to mail copies of all records of identified participants.

5.4 Clinical Examination

All study participants age five and older were requested to visit the clinic for physical examination and collection of blood and urine samples. Clinics were established in Battle Creek and Dowagiac to reduce travel for the study participants. The clinics, managed by the Borgess Medical Center, opened very early in the morning for the collection of fasting blood and urine samples and to minimize interference with work schedules of participants. At the clinic, people were asked about their current use of blood pressure drugs, and about the occurrence of new diseases since the interview. Weight, height, upper midarm circumference, and oral temperature were recorded. Blood pressure was measured with the Hawksley random-zero sphygmomanometer to avoid reader's bias (94).

Serum aliquots were prepared and refrigerated on site, and at the end of each day transported in refrigerated cases to the Borgess Medical Center for immediate processing or freezing. Frozen samples were delivered to the MDPH laboratory on a weekly basis for PCB, PBB, and pesticides testing. Tests subject to a time limit were performed within that limit; other tests were done in batches. Urine tests were performed at the clinic, except for beta-2-microglobulin, which was tested at the Laboratory of Clinical Medicine in Lansing, Michigan. Data from these various sources were put on computer tapes and diskettes by the contractor for submission to the CDC-investigator.

The clinical examination was not designed to reject, confirm, or detect any specific diseases, but to provide a general view of the current health status of the individual, utilizing routine tests. A full examination with inspection by physicians, X-rays, immunological tests, sensitive neurosensory tests, invasive function tests, biopsies, radioimmunoassays, etc. was not attempted because of the enormous cost, complexity, and time required for such an effort. Testing for

biomarkers of exposure was considered of no value because biomarkers would not remain in the body one year (or more) after exposure stopped.

5.5 Validation of Health Data

The compiled medical abstracts yielded diagnoses in all possible terminologies and ICD codes, often based on the reason for seeking medical help rather than a nosologic entity. The multiple entries and dates for the same disease, the presence of other closely related and possibly identical diseases, and the presence of disorders definitely not related to VOC-exposure (e.g. injuries) necessitated some kind of data condensation. To accomplish this, the study protocol provided for an independent medical panel: one internist and two general physicians. Following written instructions, the panel reviewed the medical abstract sheet for redundancy and the earliest date of diagnosis of a given disease. The disorders were condensed to entities based on apparent similarity of recorded diagnoses. Thus, low back pain, disc disorders in the lumbar region, sciatica, spinal arthrosis, etc., were considered manifestations of the same disease entity. It was initially expected that the panel members would not only evaluate and correct the medical records for redundancy, but would also apply their medical judgement in evaluating overt diagnostic inconsistencies noted when they reviewed the combined data from the questionnaire, medical abstract sheet, and laboratory tests. However, panel members did not adjust or correct medical diagnoses, because the study participants did not undergo a full medical examination. In a number of cases, further condensation or correction of the data was necessary prior to statistical analysis.

The goal of the health assessment was to establish *for each disease analysis separately*, one validated record per participant, accomplished by reviewing the integrated information from medical records, interview, list of medications, and clinical examination. Complete agreement of these sources was not expected, since not every disease required hospitalization or medical consultation. If a diagnosis was made, it may not have been communicated to or understood by the patient, or it could have been misfiled. Most diseases do not leave biochemical markers after successful treatment and would not be picked up by routine clinical examination. Some prescription drugs may suggest the presence of a particular disease, but most drugs have multiple indications and are not disease-specific.

Inconsistency in the health data was much worse than foreseen. Major internal inconsistencies of the questionnaire responses became evident, particularly on those questions that were repeated in one or another form. Some questions were apparently not well understood; for instance, more than 50% of the people responding to the question about "other diseases of the liver" mentioned diseases of other organs, such as lung, prostate, etc. Such inconsistencies do not pose major problems, since they can be recognized and corrected. Inconsistencies between data sources are more difficult to correct, as they may have been caused by any of a number of possible, but not always identifiable errors. Correcting such disparities requires judgmental decisions, based on evaluation by the investigator of all the data. Examples and outcomes of this validation procedure for health data are provided in the disease-specific appendices to this report.

Finally, there is the issue of data completeness. Precautions were taken to include in the search for records not only the hospitals and physicians mentioned by the interviewees or in other records, but also any other hospital in a wide area around Battle Creek and Dowagiac. For instance, a hospital in Niles, Indiana, was included, since it often served Cass County residents. Because of the absence of a quality control and assurance procedure for record retrieving and abstracting, some incompleteness of the hospital data base is presumed. It is also possible that the hospital records themselves are incomplete, as was observed for some private physicians' files. It was also found that some old hospital files were destroyed or inaccessible.

5.6 State-generated Mortality and Hospital Discharge Rates

An analysis of the uncorrected 1976-1980 mortality data for Battle Creek showed that the city had significantly higher rates for eight of the state's 10 leading causes of death (heart disease, cancer, cerebrovascular disease, chronic obstructive and related lung diseases, diabetes, chronic liver disease, suicide, and "all other"). For Calhoun County, increased rates were found for chronic lung diseases and diabetes. Subsequent evaluation of the mortality data revealed many errors in the recording and coding of the place of residence. As part of the Phase II study, the Center for Health Statistics of the MDPH conducted an extensive search of records for the period 1970-1984 looking for the causes of the errors and seeking to correct them. The results of this study and of a concurrent investigation of hospital discharge rates for counties and minor civil divisions are presented in Appendix B. Excess liver disease-related mortality was sufficiently explained by alcohol-related liver disease (recorded as secondary disease). Excess cancer mortality was confirmed for Battle Creek, but not for neighboring Emmett and Pennfield townships which include the Verona exposed cohort. There was no excess mortality for diabetes.

Chapter 6 Epidemiologic Analysis

6.1 Introduction

The analytical approach focuses on maximizing the quality of the data, evaluating dose-response relationships and the consistency of the results across multiple analytical models, in order to minimize the likelihood of spurious associations and erroneous conclusions. The basic elements in the analysis of a longitudinal study are the period of followup, the health and exposure status of individuals, the dates of first diagnosis of a disease and of the start and end of exposure, and the level of exposure. The simplest analysis is one which compares the odds or incidence rates (IRs) of a disease in the exposed and reference cohorts. In essence, a cohort-specific IR is the number of cases of a disease diagnosed in a given period of followup, divided by the number of time units that all persons in the cohort contributed during the period. Measures of comparison are the odds ratio (OR) and the relative risk or rate ratio (RR) for exposed compared with reference cohorts. Several issues in estimating and interpreting such ratios need to be considered.

6.2 Followup Time (TFU)

The longer the retrospective period of observation, the greater the possibility of incompleteness of the case-finding survey: recall deficiencies are more frequent and medical files are more likely to have been destroyed. The start of the observation period (TFU1) was January 1, 1970 or the date of moving to the study area (TIN), if that date was later. This is a compromise between the desire to go back in time as far as possible, awareness that incompleteness of data increases with time, and the financial limits of the study. The end of the followup period (TFU2) was the date of interview or the date of death. In the analysis of the association of a given disease with exposure to VOCs, the followup ended at the date a person developed the disease (TDIAG). The followup time TFU (equal to $TFU2 - TFU1$) is expressed in months to accommodate the magnitude of changes in exposure level over time. The sum of individual TFUs within a given cohort constitutes the "person-time" (person-months) of the cohort. Since $TFU2 = TDIAG$ for people with a disease of interest, TFU will be different in each disease-specific analysis.

6.3 Lag Time

Obviously, the date a disease was diagnosed is not the date the disease was irreversibly initiated. Yet, from a risk assessment viewpoint, exposure and person-time should be counted from the date a disease was initiated, not as of TDIAG. However, the onset of chronic disease is insidious; even the earliest symptoms are perceived well past the date of initiation. Using the date of initiation reduces the already small number of cases diagnosed since TIN. On the other hand, ignoring the lag time in TDIAG is not consistent with biological facts. In the analysis, TDIAG is corrected for an assumed lag time of one year (for cancer, three years) by simply moving back the date of diagnosis by one year. This added year should be viewed as an acknowledgement of the need to include the lag time in the study design (intended to become a model for an environmental health study based on risk assessment principles), rather than as an estimate of the actual lag time. For pregnancy, skin diseases, and allergies, no lag time was applied, since by the nature of these conditions the lag time is short.

6.4 Validation of Health Data

A retrospective study must rely on the quality of health information generated in the past. Such information usually is far from ideal, since medical records are not made for the purpose of studying a specific disease, and are not standardized. There are ways, however, to improve the quality of the data compiled. In this study, a multisource approach was expected to maximize quality by comparing data from at least two different sources: medical records and interview. The clinical examination provided a third source useful in the validation of hypertension and diabetes. Pharmacotherapeutic information was useful in the validation of diabetes, hypertension, hypothyroidism, stomach ulcer, and epilepsy. The introduction of seemingly redundant questions in the questionnaire was intended to identify inconsistencies in responses, and provided information leading to correction of the date of first diagnosis (TDIAG). Often, a question which asked directly about a history of a given disease yielded a negative response, while indirect questions got a positive response. In making the final decision on a person's health status, the researchers' judgement of the quality of data, and thus of the validity of a case, cannot entirely be avoided. To minimize such judgement decisions, disease-specific quality criteria (regardless of the exposure status) were developed whenever possible. These have been described in Chapter 8 and Appendices 5, 6, 7, and 8.

6.5 Eligibility of Cases for Analysis

Cases were eligible for analysis (in relation to exposure to VOCs) only if they passed the validation procedure, and if the corrected (for lag time) TDIAG was later than TIN (that is, diagnosis made after people moved to the area) and later than 1-1-1970. The reasons for the latter criterion are that the health status before 1970 was uncertain (no medical records), exposure to VOCs in the Verona area started later than 1970, and it is possible that exposure in Dowagiac and Springfield existed prior to 1970.

6.6 Univariable Analysis of Exposure-effect Associations

The conventional measures of an exposure-effect association are the odds ratio (OR) and the rate ratio or relative risk (RR). Fourfold tables provide a simple and direct way to estimate these ratios. Stratifying the analysis (strata are different levels of the risk factor) is the routine method for dealing with a confounding factor. For most diseases, the study population size of 749 was too small to allow stratification for more than one factor. Stratification into "zero", "low", and "high" exposure levels permits evaluation of dose-response relationships. The border between low and high is the median of the positive values; "low" and "high" have no toxicologic connotation. The OR and RR were calculated as follows:

	Exposed			Exposed	
	no yes			no yes	
Disease no	a	b	Cases with disease	a	b
yes	c	d	Person-months (TFU)	n	m

$$OR = (a \times d) / (b \times c)$$

$$RR = (b/m) / (a/n)$$

These fourfold tables do not apply to pregnancy outcomes. Unlike chronic disease outcomes, the pregnancy followup period does not end with the outcome, since the risk of an adverse effect of exposure continues with each subsequent pregnancy. Further, the short duration (40 weeks)

puts pregnancy outcomes in the domain of subacute rather than chronic effects. For pregnancy outcomes, the OR is based on fourfold tables in which c and d refer to the number of pregnancies with, and a and b to the number of pregnancies without the abnormal outcome of interest. No RR is calculated.

6.7 Multivariable Analysis

Multivariable regression techniques allow adequate control of confounding factors, although the number of cases may still be a limiting factor. These techniques have the disadvantage that their complexity often obscures the relation between disease and exposure. The typical analytical tool for a followup study is the proportional hazard model (41, 49, 82). Cox's proportional hazard model, worked out by Harrell as a SAS program PHGLM (41), is intended for classic cohort studies. That is, all cohort members are exposed at one point in time (TFU1), and the exposure level does not change until TFU2, the date the followup period closes. In the Battle Creek scenario, as in virtually all other situations of environmental contamination, exposure levels were not constant over time, people were not all exposed at the same time and for the same duration, and TFU1 varied for those who moved to the study area after 1-1-1970. Individual wells were contaminated by from one to seven of the seven VOCs found in the well water. Since T_1 , sample concentration C , and the change in $C = f(T)$ are VOC-specific, TAEVOC and TAEVCL (composite exposure values), cannot be defined as a simple function of time T .

A modification of the conventional proportional hazard model was developed to solve this problem. The followup period was divided into 16 segments, each with its own population. The population of period I (1970), consisted of all study participants, except those whose date of entry into the study area (TIN) was later than December 1970, and those with a disease of interest diagnosed before 1970. People who developed disease within this period were assigned to the category CASE=1; all others were CASE=0. The exposure status (TFU) for this period was maximally 12 months.

The population of period II (1971) encompassed all those of period I *plus* those who entered the study area in 1971, *minus* those who developed disease in the preceding period. People with disease were assigned to CASE=1 only if TDIAG of the disease of interest was in period II. The exposure status and TFU were again determined for this period only.

This procedure was repeated for each subsequent period, thus creating 16 period-specific populations. These populations were then pooled. Each period-specific observation of an individual was treated as an individual observation with a period-specific exposure, health status, and TFU. Harrell's PHGLM was applied to this pooled population, blocking for the period of observation.

This modification of the proportional hazard model and its application in an epidemiologic study have not been reported in the literature. The modification is based on a method described by Abbott (1), who adapted the logistic regression model to cohort studies by creating a pooled population of period-specific observations. For the current study, this concept was extended to include the estimation of period-specific exposure values, to account for exposure levels changing over time. The procedure deals effectively with the fact that people in the exposed cohort may have been unexposed for some period of time, and it facilitates the calculation of TAE for each VOC and of the composite measures TAEVOC and TAEVCL, regardless of the complexity of the relation of exposure to time-specific events.

Exposure levels for the members of the Verona exposed cohort changed even during the short period of one year. By defining the exposure as a value accumulated over the index period, it is implied that the biological effect of exposure in that period is the same as if exposure were

distributed over that period at a constant level equal to the mean TAE. This assumption, is in accordance with the general use of the average dose in toxicology. It is unavoidable because of the difficulty (for reasons of computer time and the small number of cases) of partitioning the followup period into shorter periods.

The problem of the exposure status of individuals since exposure ceased has arisen because exposure stopped well before TFU2. It is difficult to view persons in these periods as "unexposed" (TAE=0), given their exposure in previous periods. Expressing exposure as the TAE incurred in the index period only, disregarding exposure in any of the previous periods would be appropriate for subacute or acute effects (e.g., pregnancy outcomes, some skin disorders) of exposure to VOCs, in view of the short biological half-life of the chemicals.

Alternatively, exposure may be expressed as the TAE of the index period *plus* all preceding periods. Thus, an individual exposed in any earlier period remains in the exposed category. This option would be the preferred one in the case of exposure to biopersistent compounds such as PCBs and DDT, or in the case of a toxic effect with a long subclinical ("incubation" plus latency) period such as cancer. A third option is to use the peak exposure during the followup period. There is no special advantage to this since the peak level is directly derived from, and related to, TAE in either the index period or the entire period of exposure. Since no decisive arguments could be forwarded in favor of either of the first two options, both have been applied in the final analysis.

Finally, there is the problem of defining the TAE for people with a disease of interest. As explained earlier, TAE is the area under the CT-curve (shaded area in Fig. 3). The most logical estimate is TAE incurred up to TDIAG, corrected for lag time. In a model with period-specific observations, this might bias the risk estimate. Because TDIAG is rarely at the end of the period, a person with the disease would qualify for a TAE lower than that of others in the same period who had the same level of exposure, but no disease. In the extreme, this might result in a negative dose-response effect. A solution for this problem is to calculate for *everyone* the TAE from TFU1 to the start of the index period.

Multivariable analysis with the modified proportional hazard model was performed with three expressions of exposure:

- option 1: TAE incurred in index period until end of period or TDIAG
- option 2: TAE from TFU1 until end of index period or TDIAG
- option 3: TAE from TFU1 until the start of the index period

6.7.1 Multivariable Analysis of Reproductive Events

A particular analytical problem is posed by reproductive events. The likelihood of becoming pregnant (and thus the risk of miscarriage) is much lower for a woman of age 40-44 than for a woman of age 20-24. Yet, both women would contribute the same five person-years of followup time to the analysis. To further complicate the issue, a woman has some control over her probability of an adverse pregnancy outcome by applying some form of contraception. Finally, a woman's risk does not stop with the outcome of the index pregnancy. She remains at risk with each subsequent pregnancy.

Therefore, there is no rationale for using the proportional hazard model for reproductive events. Instead, a conventional logistic regression analysis was used, with the pregnancy (rather than the woman) as the unit of observation. A woman contributes as many observations as she has pregnancies. Exposure status was assessed for each pregnancy as described in Appendix H. All observations were then pooled in a single data set for use with a conventional logistic

In multivariable models with TAEVOC or TAEVCL, the variable WASH was used as a measure of the frequency and duration of water use for washing. The presence of a TAE variable in an exponential model links such a measure to the level of contamination. In DOSE models, however, linkage was achieved by multiplying WASH by CS, which is the sum of the VOC concentrations in the water sample. This resulted in WASHVOC in DOSEVOC models, and in WASHVCL (WASH times sum of VOCs including chloroform) in DOSEVCL models.

The OR was calculated for all diseases in the analysis. No RR was computed for skin disorders, reproductive events, and allergies, because of the extreme unreliability of TDIAG (for skin disorders and allergies), or because the concept of person-time of followup has little relevance to the subacute or repetitive nature of pregnancy outcomes, allergies, and skin disorders. Multivariable analysis was performed if at least 20 valid and eligible cases of a disease (TDIAG later than TIN and later than 1-1-1970) were available. However, because some people had missing values for one or more of the variables involved, the number of cases in the final multivariable equation could be less than 20.

Whatever the exposure measure or model, all analyses were also performed for the restricted population consisting of the Verona exposed and the Calhoun County reference cohort. Since exposure in the Verona area did not start until 1977, cases diagnosed before 1977 were excluded. To maintain comparability between cohorts, the Cass county reference cohorts were removed from the database as well. The rationale for this alternative analysis is that the Dowagiac and Springfield cohorts have poor exposure estimates because no $C=f(T)$ could be developed. An added advantage is that the accuracy and completeness of health data are improved, because the period of recall is shorter and the retrieval of medical records is more complete. The disadvantage is the reduction in the number of cases and, therefore, in the statistical power.

The use of so large a number of analyses was felt to be necessary because of the impossibility of solving all problems with one unanimously-agreed-upon exposure expression and analytical model. The advantage of using multiple analytical approaches is that exposure-effect associations are viewed from several angles. The analytical result gains considerably in strength.

6.9 Incidence Rate and Prevalence

Because of the nature of the study design (a cohort study) the data allow the direct estimation of incidence and prevalence rates. The incidence rate (IR) is the number of new cases per 1000 population per year, observed within a certain period (in this study 1970-1985). Since the cohorts were open for people to enter at any time, the denominator of the IR is equal to the total TFU accrued by the participants during 1970-1985. As explained earlier, TFU is the number of months between the start of the followup (TFU1) to the close of the study (TFU2), or to TDIAG. In the computation of an annual IR, a factor of 12 is necessary to convert person-months to person-years. The numerator is simply the total number of all confirmed cases, regardless of when and where they were first diagnosed, since exposure to VOCs is irrelevant to the computation of a population-based IR. This also holds for the estimation of the prevalence of a disease. The prevalence is the percentage of living people with the disease of interest in the population at a given point in time (in this study, July 1985).

Prevalence and IR are computed for several reasons: a) they give an impression of the completeness of the case-finding survey if reference data are available; b) the IR of most diseases is unknown, and this study may thus provide valuable new information; c) although the prevalences of most diseases are known from national surveys (68), such information often applies to a restricted population, e.g. people of a certain age group or the non-institutionalized population.

6.10 Evaluation of the Results of Statistical Analyses

Due to the relative rareness of noninfectious diseases, even a study of several thousand persons is considered small. Since the distribution of diseases in a small population is not uniform, differences in disease incidence between subpopulations always occur, regardless of how the subgroups are formed. Hence, it is crucial to stipulate when such a difference will be considered suggestive of an exposure-disease association. Conventionally, the p-value is used as a parameter. From a risk assessment viewpoint, however, a small p-value should not be the only consideration. An excess disease incidence with a large p-value might be significant if the study size were larger. Conversely, a small p-value does not rule out a chance effect. Other parameters used in this study are as important, though less quantifiable than the p-value. Throughout the Report, p-values were derived from a one-sided Fisher's exact test for OR and RR, and from Poisson distribution for incidence and prevalence rates. These are: a) the consistency of the results across strata, exposure expressions, and analytical approaches; b) a positive dose-response effect; and c) the biological plausibility of the observed statistical association. This issue is further discussed in Chapter 9. The final conclusion drawn from the analysis must, therefore, be based on an evaluation of many issues separately and in conjunction with each other. In this process some degree of analytical judgement must be used.

Chapter 7 Quality Control for Clinical Laboratories

by
Donald L. Phillips, PhD
CDC/CEHC

7.1 Overview

The subcontractor for all clinical laboratory measurements was the Borgess Medical Center, Kalamazoo, Michigan. There was one exception to this arrangement: urinary beta-2-microglobulin was measured by the Laboratory of Clinical Medicine, Lansing, Michigan. Each of these laboratories appointed a Quality Control Coordinator who was responsible for implementing all quality control procedures. The contractors were required to:

- Maintain the initial approved laboratory method without modification.
- Include appropriate calibrating standards in each analytical run, to be used for calculating concentrations of unknown samples in that run.
- Analyze study specimens together, and analyze them separately from specimens not part of this contract.

For all quantitative measurements, the contract laboratories were required to procure two levels of commercially available, assayed, bench-quality control materials in sufficient quantity to last for the duration of the study. The laboratories were required to maintain mean and range (if appropriate) quality control charts for the bench controls used with the study specimens, and to plot the results immediately after each analytical run was completed. If an analytical run was declared out of control (by the rules listed in each section below), the contractor was required to take the following remedial actions:

- Discontinue the analysis of specimens.
- Check for personnel and recording errors associated with quality control samples.
- Check and recalibrate instruments and other parts of the system (reagents, etc.) and verify with known control materials that the process is back in control before performing further analyses.

7.2 Serum Chemistry Profile

The contract laboratory procured two quality control (QC) materials for each of the serum chemistry profile analytes and established QC limits for them on a series of at least 20 analytical runs before any specimens from the study were analyzed. Each QC material was run in duplicate in each run. Quality control limits (95% and 99% mean and range limits) were calculated from these characterization data by the Division of Environmental Health Laboratory Sciences, CDC and provided to the contract laboratory for use in quality control. In addition, the CDC provided a blind QC serum pool. This material was also characterized, and QC limits were computed from 20 analytical runs prior to the analysis of study specimens. Vials of this material were inserted in each analytical run along with the study specimens and were indistinguishable from them by the analysts. The analytical runs for a serum analyte were declared out of control if any of the following conditions were found:

- * A single QC material mean fell outside the 99% control limits.
- * Two successive mean values fell outside the 95% control limits on the same side of the mean. The last run was then declared out of control.
- * Eight mean values in succession fell either all above or all below the mean. The last run was then declared out of control.
- * A single range value fell outside the 99% limits.
- * Two successive range values fell outside the 95% limits on the same side of the mean. The last run was then declared out of control.

Duplicate samples were submitted for analysis for approximately 10% of the study participants. These were coded with different ID numbers to blind the analyst. The results from analysis of these blind split duplicates, the bench QC materials, and the CDC blind QC pool were sent periodically on computer magnetic tapes to the CDC for QC review. Any runs found out of control, and not already detected and repeated by the contract laboratory, were brought to their attention. These runs were repeated unless the apparent out-of-control situation was due to a clerical error in reporting the QC results. Coefficients of variation were computed for all split duplicates and used to gauge the precision of the analytical measurements. This precision appeared to be good and no runs were required to be repeated as a result of poor results for split duplicates.

7.3 Hematology Profile

For the hematology profile analyses, the contract laboratory procured three levels of assayed QC pools. These could not be characterized by 20 runs prior to the analysis of study specimens, however, because they were stable only for 30 days. Several different assayed lots had to be used during the course of the study. The manufacturer's control limits (95% limits) for each lot were used to determine out-of-control runs. A run was to be declared out of control and was repeated if any of the three pools gave results outside of the 95% limits established by the manufacturer. In addition, blind split duplicates were analyzed for 10% of the study participants, as was the case for the serum analytes. These data were also transmitted to CDC on magnetic tape and the duplicates showed very good precision.

7.4 Urinary Profile

For urinary beta-2-microglobulin analyses, two commercially assayed QC materials were procured by the Laboratory of Clinical Medicine. Since the arrangements for these analyses were not worked out until shortly before the initiation of specimen collection, there was insufficient time for 20 analytical runs to establish the laboratory's own QC limits. The manufacturer's 95% limits were used, and runs were to be declared out of control if either of the QC materials gave values outside these limits. These data were transmitted to CDC on paper for QC evaluation. None of the runs reported to CDC were outside of the limits.

Analyses for urinary albumin, glucose, and blood were qualitative to semiquantitative and were performed by reading dipsticks (Chemistrips, trade mark). Two QC materials, known to be either positive or negative for these analytes, were obtained by the contract laboratory. These were used to check each bottle of dipsticks before they were used on study specimens.

7.5 Serum PCB and PBB

Serum PCB and PBB levels were determined by the Michigan Department of Public Health laboratory, Lansing, Michigan. Two QC materials were prepared for each analyte by the laboratory. Lack of time prevented the establishment of QC limits from 20 analytical runs prior to the study. However, Warning Limits and Rejection Limits were established using the data from the 27 analytical runs during the study and applied retrospectively. These limits were determined as the mean \pm 20% for the Warning Limits and the mean \pm 30% for the Rejection Limits. Assuming a maximum tolerable coefficient of variation (or relative standard deviation) of 10% for these assays, the limits correspond to two and three standard deviation limits. All analytical runs fell within both the Warning Limits and the Rejection Limits. The coefficient of variation achieved by the laboratory was less than the 10% maximum, ranging from 5.5 to 9.4% for the four QC material/analyte combinations.

Chapter 8 Results

8.1 Introduction

This study yielded a large number of health outcomes and laboratory test results. The number of these is too large to permit investigating the possible relationship between each of them and exposure to VOCs. It would also be meaningless to do this since the majority of the health outcomes and many laboratory outcomes cannot be linked to VOCs. For instance, menstrual disorders, infections, atherosclerosis and its sequelae, arthritis, injuries, iron content of serum, blood sedimentation rate, weight/height index, etc. can be ignored in the evaluation of possible health effects of exposure to VOCs. As described in Chapter 5, a substantial part of the health data collected refers to disease categories rather than nosologic entities. For most diseases, the frequency of occurrence is too low for a meaningful analysis. This report therefore presents the results of the analyses of laboratory outcomes, and of diseases which:

- * Occurred frequently enough to yield more than 10 *valid* cases; *and*
- * Are identifiable as a nosologic entity; *and*
- * Fit the toxicologic profile of VOCs, *or* for which an association with VOCs could not be precluded given the current level of understanding of the pathogenic processes.

Only a few diseases occurred frequently enough for multivariable analysis. Results of these analyses are reported in detail in appendices to this report. Other diseases were less frequent, and in these cases the analysis was limited to a univariable analysis. Given the wealth of data, it would be an unjustifiable waste of public resources not to expand the analysis to include the evaluation of the association of diseases with factors other than VOCs. However, the main objective of this study is the investigation of the potential adverse effects of VOCs. Therefore, such analysis was performed only to a limited extent even though a detailed analysis of a number of issues unrelated to VOCs may be important from a public health viewpoint. As indicated above, a disease qualified for analysis on the basis of a clearly defined set of criteria. Thus, a disease may have met one criterion, e.g., a sufficiently large number, but was not analyzed if it did not meet other criteria. It needs to be emphasized that the selection of a disease for analysis was *not* based on its seriousness, and that the results are presented in order of convenience, not significance.

8.2 PCB - PBB - and Chlorinated Pesticides

Table 8.1 summarizes the test results for PCB, PBB, and chlorinated pesticides. The presence of the pesticide DDT in almost 100% of the study population, about two decades after it was banned, emphasizes its persistence in biota. As for PCB, the prevalence rate and magnitude of serum levels are within the range reported for general population samples (83).

Only two residents from the exposed areas and four from the comparison areas were found to have serum values of heptachlor-epoxide and oxychlordan above the detection limit of two ppb. The maximum value measured was 5 ppb in the exposed cohort and 2.6 ppb in the reference. These levels, and those for the chlordan metabolite trans-nonachlor, indicate that chlordan and heptachlor do not pose a significant public health problem in the study area. Since occupation did not play a role in the formation of study cohorts, this conclusion probably holds for the general population as well. However, the number of people employed in agriculture and who participated in the study was too small to extrapolate this finding to agricultural populations or other people that are likely to be exposed to these chemicals.

Table 8.1 Summary of serum values (in ppb) of people testing positive (above detection limit) for some halogenated compounds. Values are presented as the mean, median (med), and 90th percentile (90%), discounting negative (zero ppb) test results. Testing was limited to people aged five years and older.

	PCB			PBB			DDT			Trans-nonachlor		
	Mean	Med	90%	Mean	Med	90%	Mean	Med	90%	Mean	Med	90%
Exposed cohort (exposed = having lived in the contaminated area)												
Serum level of positive tests	7.2	5.5	13.3	2.4	1.8	4.7	7.7	4.1	17.0	1.4	1.5	1.8
Number tested	218			218			218			218		
% positive tests	40.4			31.2			97.7			2.8		
Reference cohort												
Serum level of positive tests	7.4	5.9	11.6	2.1	1.5	3.4	7.7	4.5	15.1	1.4	1.4	2.6
Number tested	463			463			464			464		
% positive tests	46.0			28.7			98.7			3.0		

PBB in serum is specific to Michigan residents, since this fire retardant was accidentally introduced into the food chain in 1973 (49). The extent of the spread of the chemical in Michigan is unknown, but can be estimated from the survey conducted by Selikoff and Anderson (85). In that survey, samples were taken from the general population in six areas. The total prevalence (all people with PBB at or above 1 ppb) was 38% with little variation between areas. For Kalamazoo, the sampled area closest to Battle Creek and Dowagiac, the prevalence was 39%, and the mean PBB (mean calculated from all values, including the negative sera) was 1.3 ppb. Only 1% of the people had values above 10 ppb, and the maximum level measured was 16.8 ppb. Our findings of 30% prevalence rate, a median (of all people with a detectable level) of 1.7 ppb, a maximum value of 20.1 ppb, and 0.3% of the positive tests showing values above 10 ppb, are a very close match of the pattern described for the Kalamazoo area. These data suggest that there is no clear decline in body burdens, despite six years between these two studies.

The data compiled in the current study form the largest available PBB data set permitting an analysis of the associations with diseases. PBB is known to be persistent in biota. Yet, Kreiss *et al.* (54) described an average decrement of 1 ppb per 2 years in a population selected for its high probability of exposure. The serum levels of this group were reflected in a mean (all sera included) of 23.2 ppb; close to 20% had values over 20 ppb. Following this rate of decline, the Phase II study population should have had a mean PBB value of less than 1 ppb, which is below the detection level. This was not the case. However, an exploratory analysis did not show that any of the diseases analyzed was associated with PBB as a risk factor.

It can be concluded that VOC exposure is unrelated to the serum values of other exogenous chemicals and that the levels of PCB and DDT are similar to levels in other populations. The largest difference in the PBB levels was between Ceresco and Barron Lake (reference areas with

means of 1.5 and 4.4 ppb, respectively), and in DDT levels between Dowagiac and Springfield (means of 6.8 and 10 ppb). The differences between neighborhoods were much too small to be statistically significant, or to indicate that they are attributable to differences in lifestyle or occupation. As mentioned in Chapter 3, there were fewer people occupationally exposed to chemicals in general in the exposed cohort than in the reference cohort.

Regression analysis and simple correlation analysis showed that serum levels of PCB, PBB and DDT were highly correlated with serum cholesterol and triglycerides, and with obesity, in agreement with their high fat solubility. How serum values relate to the concentrations in fat tissue and, thus, the body burden, is not well understood, but Steinberg *et al* (90) suggest that blood lipids may interfere in the mechanisms involved in the laboratory assay. The close correlation between these chlorinated compounds and blood lipids should be seriously considered when interpreting association of these compounds with diseases such as hypertension and diabetes.

8.3 Results of Clinical Examination

The analysis of clinical test results is less complicated than the analysis of diseases in relation to exposure. First, the tests are not subject to biases associated with subjective information. Second, clinical tests follow routine procedures, subject to thorough QA/QC procedures.

The first stage of the statistical analysis consisted of comparing the prevalence of abnormal values found in the exposed and unexposed study cohorts. An abnormal value is defined as any test outcome exceeding the limit of the range of normal values in use in the laboratory that performed the tests. Since the true boundary of normal values is a range, rather than a single number, the values which exceeded the limit of the normal range were categorized into two "degrees of abnormality". Degree I comprises values which exceeded the limit by less than 10 %; Degree II comprises values 10% or more in excess of the limit. Only one test (lactose dehydrogenase or LDH) had values exceeding the upper limit by more than 100%.

Table 8.2 shows the prevalence rates for abnormal outcomes of the clinical tests, and the OR values calculated from a fourfold table for exposure (defined as having lived in the contaminated area). For the Degree I abnormal values, the table shows ORs below unity (excess of abnormalities in the reference cohort) for nearly all tests. The only OR clearly above unity is for LDH. For the Degree II abnormal values a few more elevated ORs were found, but the large majority of ORs still indicate an excess of abnormal values in the reference population. None of the tests had ORs above unity for both categories of abnormal values. These results are conclusive evidence that living in the exposed area is not associated with an adverse effect on blood chemistry. When aware that laboratory tests indicate *current* health status. It is unknown how the values relate to *past* health status. Although abnormal values may indicate a recent change, they usually reflect a disorder of longer standing. To investigate the relation between abnormal test outcomes and a better measure of exposure (TAE or DOSE), one would need an estimate of the date when the abnormal value was first present. The study design, however, did not provide for a search for historical clinical chemistry data. Hence, TDIAG is unknown and TAE or DOSE cannot be calculated properly.

In an attempt to evaluate the effects of TAE and DOSE, assuming that the laboratory values represent a *recently* developed status, TAE and DOSE were calculated as if TDIAG equalled the date that the followup of the cohorts ended (TFU2). A conventional correlation test was then done on continuous values. TAEVOC and TAEVCL were found to be negatively correlated with all tests except LDH. Direct bilirubin, hemoglobin, and beta-2-microglobulin had positive correlations with TAEVOC only. In these positive cases, the *r*-value was negligibly small ($r < 0.06$). DOSEVOC and DOSEVCL had more positive correlations, but the coefficients were also

negligibly small ($r < 0.12$). Linear regression models controlling for WASH, sex, age, alcohol use, exogenous chemicals, and occupational exposure to VOCs, showed no positive and significant ($p < 0.1$) relationship of any of the tests in Table 8.2 with TAEVOC, TAEVCL, DOSEVOC, and DOSEVCL. Odds ratios for abnormal test outcomes and overweight, calculated by fourfold tables based on dichotomous exposure. Variables showed a predominance of values below unity.

Table 8.2. Number of people with abnormal results of clinical tests. A result is abnormal if it exceeds the upper limit (for hemoglobin the lower limit) of normal values by up to 9.9% (I) or by 10% and more (II).

Category of abnormality Number (N) and % abnormal	Prevalence in exposed Area				Prevalence in reference Area				Odds Ratio	
	I		II		I		II		I	II
	N	%	N	%	N	%	N	%		
Iron	9	4.0	12	5.4	22	4.8	15	3.3	0.85	1.67
Iron binding capacity	12	5.4	8	3.6	41	8.9	10	2.2	0.59*	1.60
Hemoglobin (Hb)	4	1.8	1	0.5	23	5.0	4	0.9	0.34*	0.49
Hemoglobin A1C (HbA1C)	2	0.9	5	2.3	8	1.7	5	1.1	0.51	2.06
Glucose	7	3.2	9	4.1	22	4.7	21	4.5	0.64	0.86
Cholesterol	9	4.1	4	1.8	20	4.3	12	2.6	0.91	0.68
Triglycerides	6	2.7	36	16.2	24	5.2	88	18.9	0.48*	0.78
Body mass index (QI)	29	13.0	65	29.1	86	18.4	143	30.6	0.60*	0.81
Blood urea nitrogen (BUN)	1	0.5	5	2.3	3	0.6	21	4.5	0.67	0.48*
Creatinine	0	0	0	0	3	0.6	3	0.6	0	0
Uric acid	4	1.8	2	0.9	8	1.7	5	1.1	1.03	0.82
Beta-2-microglobulin	1	0.5	15	6.8	2	0.4	20	4.3	1.06	1.58
Direct bilirubin	0	0	1	0.4	0	0	0	0		
Total bilirubin	12	5.4	8	3.6	41	8.9	10	2.2	0.59*	1.60
Serum glutamic oxalo- acetic transaminase (SGOT)	11	5.0	10	4.5	27	5.8	27	5.8	0.81	0.74
Serum glutamic pyruvic transaminase (SGPT)	1	0.5	11	5.0	10	2.2	28	6.0	0.20*	0.78
Gamma glutamyl trans- peptidase (GGTP)	3	1.4	14	6.3	10	2.2	47	10.1	0.58	0.58*
Lactate dehydrogenase (LDH)	16	7.2	10	4.5	20	4.3	27	5.8	1.67*	0.77
Alkaline phosphatase	26	11.6	37	16.5	60	13.0	109	23.7	0.78	0.61*

- 1) Limits of normal values were adopted from the Borgess Medical Center (the contract laboratory). Overweight is expressed in the body mass or Quetelet Index (QI), which is weight (kg)/height² (m). The upper normal limit of QI is 25, which is the median of the values in the study population.
- 2) Samples were taken from 224 people in the exposed and 560 in the reference areas. For some tests, these numbers are somewhat smaller due to a sample deficiency (hemolysis, breakage, insufficient volume).
- 3) The asterisk (*) indicates $p < 0.1$ (one-tailed Fisher's exact test).

In summary, although the value of these analyses is dubious because of the impossibility of finding a proper TDIAG, the results are in line with the earlier conclusion, that exposure to VOCs has not adversely affected health, as measured by the clinical examination. Surprisingly, the data suggest an excess of abnormal values in the reference cohorts. It should be kept in mind that whenever a large number of comparisons are made, statistically significant correlations are bound to occur merely by chance. However, there are too many negative associations of the clinical test results with exposure and too many statistically significant low ORs to be explained by chance alone. On the other hand, to infer a "protective" effect of VOCs is not supported by the known biologic effects of VOCs. The presence of an unidentified risk factor(s) in the reference cohort, especially the Battle Creek City reference cohort, is another possible explanation.

The extremely tight QA/QC controls for laboratory procedures rule out the possibility that abnormal test values in the reference areas can be ascribed to laboratory errors. The high prevalence of abnormal values may also mean that the upper limit of normal values, in use at the hospital of the subcontractor may have been too low for at least some of the clinical tests. The prevalence in the study population, regardless of exposure and category of abnormality, is the sum of abnormal tests (N) (from the four columns in Table 8.2) divided by 684 (total number of tests done). Upper normal limits are assumed to be the 95th percentile of the values in apparently healthy people. A prevalence of 5% or less of the 684 tests was found only for Hb, HbA1C, BUN, creatinine, uric acid, and direct bilirubin. A prevalence of more than 10% among the 684 people tested was found for the iron binding capacity (10.4%), total bilirubin (10.4%), SGOT (11.0%), GGTP (10.8%), LDH (10.7%), alkaline phosphatase (33.9%), and triglycerides (22.5%). Naturally, the study population did include a number of people with an illness which may give rise to abnormal values. At present, it is unknown how large that number is, but it is unlikely to be so large that it could explain all of the excess of abnormal values, in particular the large number of abnormal triglycerides and alkaline phosphatase values.

8.4 Diabetes Mellitus

A full report on the analysis of diabetes is given in Appendix E. This section summarizes the results. The study yielded 43 reports from various sources that diabetes might be present. After data validation, 28 cases remained, of which 20 were eligible for analysis. The prevalence rate was 3.4% of the total population alive by July 1985, 4.7% of men and 2.3% of women. The prevalence among men was comparable to the NCHS rate, but the rate for women age 20-74 was half of that found for the same age group by the National Center for Health Statistics (NCHS) in the National Health and Nutrition Examination Survey (NHANES II) of 1976-1980 (70). No explanation has been found for this high male/female ratio. The incidence rate (IR) for 1970-1985 was 2.30/1000 total population/yr. This is probably the first measured population-based IR for all types of diabetes in whites of all ages and both sexes in the U.S.A.

The ORs and RRs showed an excess of diabetes in the exposed cohort if exposure was defined as TAEVOC, or DOSEVOC, although the dose-response relationship was negative. ORs and RRs were below unity if the measures of exposure were TAEVCL and DOSEVCL. This inconsistency suggests that the increased ratios for TAEVOC and DOSEVOC were *not* related to a toxic effect of VOCs. Limiting the analysis to the Verona exposed and the Calhoun reference cohort (for better exposure data) did not change this conclusion.

Multivariable analysis included the variables WATER, WASH, and age at start of followup. The small number of cases prohibited the use of more covariates. The coefficient for the exposure variable was statistically insignificant, and WATER and WASH sometimes had a negative coefficient. In summary, the combined results show that there is no evidence of a toxic effect

of exposure to VOCs, at the levels prevailing in the drinking water in the study areas, with regard to diabetes.

Significant risk factors emerging from the diabetes data were: a positive family history of diabetes, obesity, and being a male. The latter was unexpected and may be a chance effect despite the low p-value. Other positive risk factors, with less impact on the magnitude of the risk, were age, education level, SGPT and GGTP (liver enzymes), and triglycerides. *Negative* associations with a significant p-value were found for DDT, urinary beta-2-microglobulin (indicator of renal function), and occupational exposure to chemicals predominantly composed of VOCs (WORK1, Chapter 6). The latter strengthens the conclusion that there is no observed association of VOCs with diabetes. A positive family history was associated with an OR (computed from a model controlling for age and gender) of 4.1 (95% confidence limits 1.80-9.34), a value that could change slightly if the model controlled for other significant risk factors.

8.5 Hypertension

A full report on the association of exposure to VOCs and risk of hypertension is given in Appendix F. This section summarizes the results. There were 174 people with some indication from various sources that hypertension might be present. After validation of the data, 115 cases remained, representing a prevalence of 15.5% of the total population alive by July 1985 (20.1% of the population age 15 and older). The IR for 1970-85 was 9.69/1000/yr for all ages (11.41/1000/yr for the population age 15 and older), with no significant sex difference. The prevalence in the population age 18-74 is 1.5 times higher than the NCHS estimate for the U.S. population (68).

Eighty four cases were eligible for analysis. The ORs and RRs from fourfold tables were well below unity with a consistently negative dose-response effect across populations and exposure expressions. The deficit of hypertensive cases in the exposed cohort was statistically significant. A negative association was also found in multivariable models with the covariates WATER, WASH, age at start of followup, sex, the body-mass index, and renal function. A negative association of VOCs with systolic and diastolic blood pressure was shown by correlation tests and multilinear regression analysis. In summary, there is unequivocal evidence that exposure to VOCs, (at the levels which prevailed in the drinking water) had no adverse effect on blood pressure or the risk of hypertension. The cause of the excess of cases in the reference cohort is unknown.

Further analysis of the data showed that obesity, education level, serum levels of uric acid, triglycerides, DDT, and occupational exposure to VOCs (WORK1, Chapter 6) were factors positively associated with the risk of hypertension. In view of the above negative associations it is possible that the positive association between hypertension and *occupational* exposure to VOCs occurred by chance (multiple-comparison effect). A substantial proportion of the hypertensives were not aware of the disorder. In some cases there was indication of over-treatment with antihypertensive drugs, resulting in some people with a clear *hypotension*.

8.6 Gall Bladder Disease

A full report on the association of exposure to VOCs and risk of gall bladder disease is given in Appendix G, of which this section is a summary. In the analysis, gall bladder disease included cholelithiasis (gall stones, more than 90% of the cases) and chronic cholecystitis (inflammation). Of the 55 people with some indication that gall bladder disease might be present, 48 cases were validated, representing a prevalence of 6.3% of the total population alive by July 1985 (men 3.2%, women 8.9%). The prevalence dropped to 4.5% (men 2.6%, women 6.1%) if the

case definition was restricted to cases with a positive medical record. The IR for the study period 1970-85 (using the broader case definition) is 3.89/1000 population (all ages)/yr (men 1.87, women 5.67). The sex difference is statistically significant ($p < 0.05$). Whatever the case definition, the prevalence is about 10 times higher than observed by the NCHS in the US population (69). Although the NCHS survey was probably incomplete, this would not be sufficient to explain the very high prevalence in the study area. No other reports on the IR for gall bladder disease have been found in the literature.

There were 28 cases eligible for analysis, yielding ORs and RRs which were less than unity (statistically significant for TAEVCL and DOSEVCL) across exposure expressions, case definitions, and populations. When exposure was defined as "having lived in the contaminated area", the OR was above unity, albeit with a large p-value. However, this definition of exposure has no scientific value; it ignores actual exposure and the proper time-sequence of exposure and disease.

The multivariable analysis included WASH, WATER, sex, and age at the start of the followup period as covariates. The number of cases was too small to permit analysis with more covariates. The coefficients for the exposure variable were consistently negative across exposure and case definitions, except for a positive coefficient with ($p > 0.8$) for TAEVOC and the limited population of the Verona exposed and Calhoun reference cohorts. These findings are in line with the results from the fourfold tables, and are sufficient evidence that exposure to VOCs in drinking water had not increased the risk of developing gall bladder disease. The excess of cases in the reference cohort remains to be explained.

Further analysis showed that age, obesity, being a female, and having diabetes were positively associated with the risk of gall bladder disease, while blood urea nitrogen was negatively associated.

8.7 Abnormal Pregnancy Outcomes Other Than Birth Defects

A full report on abnormal pregnancy outcomes other than birth defects is given in Appendix H. This section summarizes the results of the analyses of miscarriage, prematurity, and low birth weight. The study yielded data on the reproductive history of 312 women, 249 of whom had at least one pregnancy. In total, these women had 808 pregnancies, 11% of which ended in miscarriage, 2% in stillbirth, 4% in premature births, and 7% in low birth weight newborns. The prevalence of pregnancy loss (miscarriage and stillbirth) in the study female population age 15-44 is 28.9%, slightly higher than was reported for the USA population by the NCHS. The difference was not statistically significant.

The statistical analysis of the association of abnormal pregnancy outcomes with exposure was based on the pregnancy (rather than the woman) as the unit of observation. Of the 808 pregnancies, 199 were eligible for analysis, including 29 abnormal pregnancy outcomes. Univariable analysis yielded ORs lower than unity, indicating a deficit of cases in the exposed population, for all outcomes except prematurity. With regard to prematurity, the OR was zero (no cases among the exposed) for TAEVOC and DOSEVOC, and slightly above unity for TAEVCL and DOSEVCL. This inconsistency is almost certainly due to the very small number of eligible cases of prematurity ($N = 7$). None of the ORs reached a significance level below $p = 0.2$.

Multivariable analysis was done for miscarriage, pregnancy loss, and "any abnormal event". The analysis, which controlled for age at the index pregnancy, WATER, and WASH, yielded negative coefficients for the exposure variable for all outcomes; this is consistent with an OR lower than unity and a negative dose-response effect. Although the very small number of cases did not justify a multivariable analysis for prematurity and low birth weight, an exploratory use of the analytical models again showed negative coefficients. Accordingly, it was concluded that

exposure to VOCs, at the levels prevailing in the drinking water, has had no adverse effect on pregnancy outcomes.

8.8 Congenital Defects

VOCs have not been shown to be teratogenic in animal experiments, although they are definitely fetotoxic at high doses. Fetotoxic effects, such as low birth weight, miscarriage, and stillbirth, are addressed above and in Appendix H. This section focuses on the analysis of congenital defects, a term describing defects present at birth (but not necessarily diagnosed at that time) and associated with a disturbance in the fetal development. The defects included in the analysis are listed below, but the interviews yielded 19 additional "defects" ranging from cerebral palsy to "failure to thrive". There were 28 valid cases, of which 11 were eligible for analysis.

Table 8.3 Congenital defects (28 valid cases) according to type.

Skeleton	:	Sternum (N=2), limbs (N=9), achondroplasia (N=1)
Neural tube	:	Anencephalia (N=1), spina bifida (N=1)
Cardiovascular	:	Septal defects (N=4), vascular branch (N=1), unspecified (N=4)
Urogenital	:	Hypospadias (N=2), malformed ureteropelvic junction (N=1)
Oral	:	Cleft palate (N=2)

Data validation was possible only to a limited extent, since seven cases reported by interviewees referred to events before 1970 (a period for which no medical records were compiled), and no medical record was found of five cases born since 1-1-1970. These five unconfirmed interview cases consisted of four heart defects (three cases of "murmur", and one described as "heart problem") and one "pigeon toe", a condition which could have been acquired. These five unconfirmed and dubious cases were not included in the analysis. Some obvious defects were not mentioned by the parents: two cases of cleft palate (one also had a cleft lip), a case of hypospadias, and a case of anencephaly. These cases were found in medical records, but were not mentioned in interviews.

In the analysis, the pregnancy and not the woman was the unit of observation, because the probability of an adverse outcome is directly related to the number of pregnancies. Pregnancies ending in miscarriage were not included in the analysis. The exposure values were calculated for each pregnancy separately (see Chapter 6 and Appendix H). In the analysis, exposure values were either dichotomous (yes/no), or categorized as zero - low - high, using the rounded median of positive values to separate low from high. Table 8.4 shows the ORs, computed as described in Chapter 6, from fourfold tables. The ORs are decreased or increased depending on the exposure expression, but the p-values are very large, and the dose-response relationship is consistently negative. This was also observed when the analysis was limited to the Verona exposed cohort and the Calhoun reference cohort.

A simple logistic regression with DOSEVOC or DOSEVCL, using continuous values, confirmed the negative dose-response relationship. The number of cases was too small for multivariable analysis. It can be concluded from the combined results that there is no evidence that exposure to VOCs (at the levels which were present in the drinking water) is associated with the risk of a

congenital defect. This agrees with the negative findings for other pregnancy outcomes. The results were based on an analysis of pooled data and there is no information indicating whether pooling of different kinds of defects may have distorted the results. However, given the very small number of cases, an analysis of pooled data was the only option. A larger study size would probably not have changed the outcome of the analysis, given the consistently negative dose-response effect and decreased ORs. A study size larger by several orders of magnitude would be required if each type of congenital anomaly were to be analyzed separately.

Table 8.4 Odds ratio (OR) of birth defects. The unit of observation is any pregnancy not ending in miscarriage. The p-values were derived from a one-sided Fisher's exact test.

Exposure expression is TAE or having lived in the contaminated AREA														
Defect	Contam. AREA		Exposure = TAEVOC (VOCs)				Exposure = TAEVCL (VOC+Chl)							
	No	Yes	Low		High		Low		High					
			Exposed	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed				
No	163	78	135	30	135	14	135	16	79	86	79	64	79	22
Yes	10	4	10	1	10	1	10	0	4	7	4	6	4	1
OR =	0.84		0.45		0.96		0		1.61		1.85		0.90	
p =	0.51		0.39		0.67		0.34		0.34		0.27		0.70	

Exposure expression is DOSE												
Defect			Exposure = DOSEVOC (VOCs)				Exposure = DOSEVCL (VOC+Chl)					
			Low		High		Low		High			
	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed	Exposed		
No	138	26	138	18	229	12	82	82	82	66	82	16
Yes	10	0	10	1	10	0	4	6	4	6	4	0
OR =	0		1.27		0		1.50		1.86		0	
p =	0.19		0.58		0.61		0.39		0.27		0.50	

Two studies have associated congenital anomalies with drinking contaminated water. Dorsch and coworkers (22) found an excess of neural tube and musculo-skeletal defects in the offspring of women exposed to drinking water containing more than 5 ppb of nitrate. The water source was not tested for other chemicals, however, nor was the daily amount of water consumed taken into account. In the Woburn study (55), an excess of eye and ear anomalies was attributed to water contaminated with a mixture of chemicals, among which were some of the VOCs found in the current study. In the Battle Creek Health Study, there was not a single case of ear anomaly, and the only eye disorders were two cases of strabismus (one unconfirmed by a medical record) and one case of "lazy eye", neither of which is a "defect" in the sense of disturbed fetal development.

Birth "defect" is a term with many meanings. Some investigators have included any disorder in early childhood, such as pyloric hypertrophy. Others have included disorders probably related

to labor and delivery, such as cerebral palsy. This complicates comparisons on the basis of IR or prevalence rates which include this disorder. Considering this and the limited possibility for case validation, no IR or prevalence were computed. Bierman *et al* found that 7.6% of 1963 live births had a congenital defect (9). In Bierman's population-based prospective study, disorders such as low birth weight, undescended testicle, and umbilical hernia were included. In the Battle Creek study, the prevalence of 3.5% defects with a restricted case definition (28 cases per 808 live births) should therefore be seen as a lower bound estimate if compared with findings from other studies. The value of 3.5% suggests that the frequency of congenital defects in the study area does not differ markedly from that observed elsewhere, and it certainly is not higher.

8.9 Thyroid Disease

Many chemicals may act as goitrogens or affect thyroid function (47). This is not known for VOCs, but a test of thyroid function is not a routine component of the test battery in animal or human studies so there are no animal or human data. In total, 33 people had some information from interview, medical record, or drug prescriptions which was suggestive of thyroid disease. The interview did not directly address thyroid disorders, but some people volunteered that they had a thyroid problem, when responding to general questions on health status. Seven persons without a positive medical record used thyroid supplements. The 33 reports included two reports of hyperthyroidism; 20 of hypothyroidism; two of goiter (one identified as adenoma), three of thyroiditis; and six undefined cases. The latter included four interview reports of "thyroid disease", which were unsupported by medication or medical record, and two cases of a physician's record of "thyroid problem?". These subjects did not report a thyroid problem and did not use thyroid supplements. The analysis has focused on hypothyroidism because other disorders were too rare for a meaningful analysis. A case of hypothyroidism was valid if the medical record was positive or if the interview was positive *and* supported by at least the use of thyroid supplements.

The 20 valid cases of hypothyroidism represent a prevalence rate as of July 1985 of 3/341 or 0.88% for men, and 17/395 or 4.30% for women. If the four positive interviews (all with women) unsupported by a medical record or use of thyroid drugs were included, the prevalence rate for females would increase to 5.32%. The National Health Survey (69) did not report on hypothyroidism, but mentioned a prevalence of "all thyroid conditions" (men 0.40%, women 2.20%) and "diseases other than goiter" (men 0.21%, women 1.14%). The prevalence of all (validated) conditions in the current study (comprising fewer diseases than the NCHS category) would be 5/341 = 1.47% (men) and 24/395 = 6.08% (women). Thus, the rates found in the current study seem to be a multiple of the estimates for the U.S. population. No explanation has been found. The IR for the period 1970-1985, computed as described in Chapter 6, equals 16/125411 person-months or 1.53/1000/yr (Poisson 95% confidence limits 0.88-2.49). The sex-specific IRs are 0.62 cases/1000 men/yr (1.3-1.81), and 2.31 /1000 women/yr (1.23-3.96). No reference data on the IR were found.

Of the 20 validated cases of hypothyroidism, only 12 were eligible for analysis (TDIAG later than 1-1-1970 and TIN), too few for multivariable analysis. Table 8.5 shows the results of an analysis using fourfold tables with OR and RR as the measure of association, as described in Chapter 6. To enable an evaluation of the dose-response relationship, exposure levels were categorized as zero - low - high. A fourfold table for exposure (expressed as having lived in the exposed areas), is also presented to demonstrate the effect on the OR of an improper exposure expression (one which ignores the temporal cause-effect sequence and fails to ascertain the actual exposure status). Since no eligibility criteria for this exposure expression applied other than TDIAG > TIN, the number of cases is larger than in the other fourfold tables.

Table 8.5. Odds ratio (OR) and relative risk (RR) of hypothyroidism; p-values were derived from a one-sided Fisher's exact test.

Exposure expression is TAE or having lived in the contaminated AREA								
Hypothyroid	Contam. AREA		Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)		
	No	Yes	Exposed No	Exposed Yes	High Exposed No	High Exposed Yes	Low Exposed No	Low Exposed Yes
No	484	245	512	217	512	104	512	113
Yes	11	5	8	4	8	4	8	0
OR =	0.90		1.18		2.46		0	
p =	0.54		0.50		0.14		0.21	
RR =			1.24		2.81		0	
p =			0.47		0.09		0.20	

Exposure expression is DOSE							
Hypothyroid			Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL(VOC+Chl)	
	Exposed No	Exposed Yes	Low Exposed No	Low Exposed Yes	High Exposed No	High Exposed Yes	High Exposed Yes
No	527	189	527	90	527	99	265
Yes	8	3	8	3	8	0	6
OR =	1.05		2.20		0		0.49
p =	0.59		0.21		0.26		0.19
RR =	1.18		2.46		0		0.48
p =	0.50		0.14		0.21		0.16

The ORs and RRs were slightly above unity for exposure expressed as TAEVOC or DOSEVOC, and were far below unity for TAEVCL and DOSEVCL, without statistical significance. Whatever the exposure expression, the dose-response relationship was strongly negative, and there was not a single case in the higher exposure group. Similar results were obtained from an analysis limited (for reasons of better exposure data, Chapter 4) to the Verona exposed and the Calhoun County reference cohorts. In summary, the data show that exposure to VOCs, at the levels which prevailed in the drinking water, was not a risk factor for hypothyroidism. A larger study size would probably not have changed this conclusion, given the strongly negative dose-response effect.

8.10 Epilepsy

In this section, the term epilepsy encompasses epilepsy, seizures, or convulsions, conditions specifically mentioned in the questionnaire. VOCs are known to affect the central nervous system. However, there have been no reports relating to epilepsy or convulsions. Of the 19

people with possible epilepsy, 16 had a positive medical record. In five cases, the disorder was secondary to a cerebrovascular accident, brain tumor, or meningitis. Except for these five cases, all cases were considered valid because they were confirmed by a positive chemical record, or a positive interview supported by specific therapy.

Table 8.6. Odds ratio (OR) and relative risk (RR) of epilepsy; p-values were derived from a one-sided Fisher's exact test.

Exposure expression is TAE or having lived in the contaminated AREA								
Epilepsy	Contam. AREA No Yes	Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)			
		Low		High	Low		High	
		Exposed No Yes	Exposed No Yes	Exposed No Yes	Exposed No Yes	Exposed No Yes	Exposed No Yes	
No	488 247	515 220	515 107	515 113	241 494	241 376	241 118	
Yes	5 2	5 1	5 0	5 1	1 5	1 4	1 1	
OR =	0.99	0.47	0	0.91	2.44	2.56	2.04	
RR =		0.49	0	0.89	2.83	3.05	2.19	
p =		0.45	0.44	0.69	0.30	0.28	0.30	

Exposure expression is DOSE								
Epilepsy		Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)			
		Low		High	Low		High	
		Exposed No Yes	Exposed No Yes	Exposed No Yes	Exposed No Yes	Exposed No Yes	Exposed No Yes	
No		530 191	530 93	530 98	271 450	271 347	265 105	
Yes		5 1	5 0	5 1	1 5	1 4	6 0	
OR =		0.56	0	1.48	3.01	3.12	2.63	
p =		0.50	0.45	0.64	0.28	0.28	0.48	
RR =		0.61	0	1.10	3.52	3.73	2.88	
p =		0.54	0.50	0.63	0.21	0.21	0.45	

The 14 valid cases, of which 13 were diagnosed in the period 1970-1985, yielded an IR (for 1970-1985) of 13/125562 person-months or 1.24/1000/yr (Poisson 95% confidence limits 0.66-2.12). No reference data on the IR were found. The prevalence in the total population alive by July 1985 was 14/736 or 1.90% (1.04-3.19%). There was no significant difference in the gender-specific IR or prevalence. The prevalence was about one half of the NCHS rate of 3.7% for the U.S. population (69). However, it should be noted that the NCHS data were derived from a household interview (one respondent per household) with no validation procedures in place.

Only 6 of the 14 cases were eligible for analysis (TDIAG later than 1-1-1970 and later than TIN). The results of the univariable analysis with TAE and DOSE as described in Chapter 6 are shown in Table 8.6. Increased ORs and RRs were observed for exposure defined as TAEVCL or DOSEVCL, but the dose-response relationship was uniformly negative. It should be clear,

however, that the number of cases is too small for other than a negative conclusion. The study did not show evidence of a positive association between exposure to VOCs and epilepsy.

8.11 Ulcer

The interview did not directly address ulcer, but 17 persons provided information, suggestive of peptic ulcer, in response to general questions on current and past health experience. The data from medical records and interviews, and information on the use of antacids and/or cimetidine (suppressor of gastric acid and pepsin secretion), indicated 35 people with potential cases of ulcer. After validating the data, these people were categorized as ULCER=1 (medical record positive), ULCER=2 (interview positive, use of medication), and ULCER=3 (no medication, no medical record, interview positive). For the analysis, only ULCER=1 and 2 were accepted as valid cases. The use of antacids in the face of negative interview and medical record was not considered to be evidence of ulcer. No distinction was made between peptic, stomach, duodenal, or bleeding ulcer, in view of the small numbers and because the medical records often mentioned more than one of these categories for the same patient. Of 27 valid cases, 23 were eligible for analysis (TDIAG later than 1-1-1970 and later than TIN).

Table 8.7 Odds ratio (OR) and relative risk (RR) of peptic ulcer, defined as ULCER = 1 or 2. The p-values were derived from a one-sided Fisher's exact test.

Exposure expression is TAE or having lived in the contaminated AREA														
Ulcer	Contam. AREA		Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)								
	No	Yes	Low		High	Low		High						
			Exposed	Not Exposed	Exposed	Exposed	Not Exposed	Exposed						
No	479	243	507	215	507	108	507	107	237	485	237	373	237	112
Yes	16	8	17	6	17	3	17	3	9	14	9	11	9	3
OR =	0.99		0.83		0.84		0.91		0.76		0.78		0.71	
RR =			0.88		0.95		0.82		0.89		0.93		0.77	
p =			0.50		0.61		0.51		0.47		0.52		0.49	

Exposure expression is DOSE												
Ulcer	Exposure = DOSEVOC (VOCs)		Exposure = DOSEVCL (VOC+Chl)									
	Exposed	Not Exposed	Low	High	Exposed	Not Exposed	Exposed					
			Exposed	Not Exposed				Exposed	Not Exposed			
No	522	187	522	94	522	93	267	442	267	343	267	99
Yes	17	5	17	3	17	2	9	13	9	11	9	2
OR =	0.82		0.98		0.66		0.87		0.95		0.60	
p =	0.46		0.45		0.44		0.46		0.54		0.40	
RR =	0.90		1.13		0.69		1.02		1.13		0.67	
p =	0.54		0.52		0.46		0.57		0.48		0.46	

A univariable analysis was performed using the exposure characterization and measures of association described in Chapter 6. The results are shown in Table 8.7. Virtually all ORs and RRs were below unity, indicating an excess of cases in the reference cohort. Moreover, the ratios show a consistently negative dose-response effect. Restricting the analysis to the Verona exposed and the Calhoun reference cohort (a total of 16 cases), or to cases with an unequivocally positive medical record (18 cases of $ULCER=1$), yielded similar results. These findings are sufficient to conclude that exposure to VOCs, at the levels which prevailed in the drinking water, can be ruled out as the cause of ulcer disease. Consistency in the data suggests that this outcome was not caused by too small a sample size. No attempt was made to perform a multivariable analysis in view of the heterogeneity of cases and the decreased ORs and RRs.

The prevalence of ulcer (defined as $ULCER=1$ or 2) is $25/736$ or 3.40% (Poisson 95% confidence limits 2.20-4.90) of the population alive by July 1985. Men had a slightly higher rate than women, but the difference was not statistically significant. The NCHS found a prevalence rate in the 1979-1981 survey in the general population of 1.71% for males and 1.74% for females (69). As mentioned earlier, the NCHS data originated from unvalidated household interviews, which may explain the higher prevalence in the study population. Given the higher prevalence rate, the Battle Creek study is not likely to have suffered from incompleteness of the case-finding survey. The IR for the period 1970-1985, calculated as described in Chapter 6, is $27/125258$ person-months or $2.59/1000/yr$ (1.70-3.76). The IR for males ($3.13/1000/yr$) was higher than that for females ($2.13/1000/yr$), but the difference was not statistically significant.

8.12 Kidney Disease

High doses of VOCs may be toxic to the kidneys (Chapter 3). The interview therefore contained direct questions on kidney stones, "nephritis", renal failure, urinary tract infection (UTI), hematuria, proteinuria, and "other" disorders in addition to the general questions on the medical history. The compiled data suggested that 34 people might have had a kidney disease other than UTI: four renal failure, 16 renal stones, nine pyelonephritis or "nephritis", and five other disorders. Cases were considered valid if they were supported by a medical record or laboratory outcome. Positive interview responses without such support, UTI, and cases of a clearly inflammatory origin were ignored. There were 18 valid renal disease cases, 11 of which were eligible for analysis (TDLAG later than 1-1-1970 and later than TIN). The largest group was seven eligible cases of renal/ureteral stones.

The prevalence of kidney (or ureteral) stones in the population alive by July 1985 was $10/341 = 2.93\%$ (Poisson 95% confidence limits 1.41-5.39) of the male population, and $4/395 = 1.01\%$ (0.28-2.59) of the women. If only medical record cases were counted, the prevalence among women dropped to 0.25% , while the rate for men remained unchanged. The NCHS estimate of the prevalence in the U.S.A. population for the period 1979-1981 is 0.43% for men and 0.31 for women (69). Regardless of the case definition, the prevalence rate among men in the current study remains a multiple of that found by the NCHS. The NCHS data were based on household interviews and this may explain in part the higher rate in the current study. The IR for the period 1970-85 for medically confirmed cases of kidney stones is $2.10/1000$ men/yr (1.01-3.85), and $0.18/1000$ women/yr (0.01-0.98). No reference data on the IR were found.

The number of cases eligible for analysis was too small for a multivariable analysis. Using ORs and RRs as measures of exposure-disease associations, as described in Chapter 6, a univariable analysis was carried out. The results are shown in Table 8.8. Regardless of the exposure expression, the ORs and RRs were less than unity, with a negative dose-response relationship. This observation was not changed when the analysis was limited to the Verona exposed and the Calhoun County reference cohorts. Similar results were found if the analysis focused on kidney stones plus validated cases of "other" renal disorders. At any rate, the number of cases is too small for a conclusion other than that there is no evidence that exposure

to VOCs, at the prevailing levels, is associated with an increased risk of renal disorders. The excess of cases in the reference cohort, if true, remains to be explained.

Table 8.8. Odds ratio (OR) and relative risk (RR) of renal or ureteral stones. The p-values were derived from a one-sided Fisher's exact test.

Exposure expression is TAE or having lived in the contaminated AREA													
Stones	Contam. AREA		Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)							
	No	Yes	Exposed	Low	High	Exposed	Low	High	Exposed				
				Exposed	Exposed		Exposed	Exposed		Exposed			
No	471	247	519	219	No exposed cases			241	497	241	380	241	117
Yes	7	1	7	0				3	4	3	4	3	0
OR =	0.28		0			0.65		0.85		0			
p =	0.20		0.09			0.42		0.56		0.31			
RR =			0			0.75		1.01		0			
p =			0.10			0.49		0.65		0.33			

Exposure expression is DOSE													
Stones			Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)							
	No	Yes	Exposed	Low	High	Exposed	Low	High	Exposed				
				Exposed	Exposed		Exposed	Exposed		Exposed			
No			534	190	No exposed cases			271	453	271	350	271	103
Yes			17	5				3	4	3	4	3	0
OR =			0			0.80		1.03		0			
p =			0.12			0.53		0.64		0.38			
RR =			0			0.94		1.23		0			
p =			0.14			0.62		0.54		0.41			

8.13 Cancer

At least some of the VOCs are considered to be carcinogenic to laboratory animals, although this is mostly based on dubious evidence (see Chapter 3). The main problem in the analysis of cancer as a single health outcome is that cancer is not a single disease, while epidemiologic characteristics are type-specific. Pooling all cancers into one database to increase the number of cases forces the use of implicit and explicit assumptions and admittedly stretches the analysis beyond the limits of what is scientifically justified. Inferences from these results may therefore not be valid. However, in communications with exposed residents, cancer ranked as the disease of greatest concern. This warrants a report in more detail than those presented for other diseases, but it must be recognized that a detailed analysis is not necessarily a better or a more satisfactory analysis.

The data sources were interviews, medical records, clinical chemistry test results, and a list of prescribed drugs. Clinical tests relevant to cancer, including a blood cell count, blood cell differential, and acid phosphatase (males age 50+), did not yield results indicative of cancer. The list of prescribed drugs did contain cytotoxic drugs, but there were no cases not already identified otherwise. The combined data sources yielded the following 46 reports of possible cancer:

Table 8.9 Reports of possible cancer (all sources).

9	Benign lesions: sinusitis (!), lump in breast, suspect mammogram, colon adenoma, carcinoid in appendix, cervical dysplasia, uterine tumor (medical record negative), subcutaneous nodular fibrosis
3	Premalignant (in-situ) lesions (2 cervix and 1 larynx, all confirmed)
10	Skin cancer (9 confirmed: all basal cell carcinoma)
5	Leukemia (3 confirmed)
5	Breast cancer (all confirmed)
3	Cancer of corpus or cervix uteri (all confirmed)
2	Colon cancer (both confirmed)
2	Lung (both confirmed)
2	Metastatic cancer, primary organ unknown (both confirmed)
5	Single cases of lymphoma, malignant craniopharyngioma, and carcinoma of tongue, kidney, and prostate (all confirmed)

In the list above, "confirmed" indicates that a positive medical record was found. In the validation process, these 46 reports were categorized according to the indications of malignancy. The results were as follows:

- CA = 0: The lesion reported was benign (9 cases).
- CA = 1: A positive medical record of invasive cancer (31 cases).
- CA = 2: A positive interview response not confirmed by a positive medical record; none of the subjects reported hospitalization (3 cases).
- CA = 3: The medical record mentioned in-situ cancer (3 cases).

Of the 34 people with a positive medical record, five (15%) responded negatively when asked for a history of cancer (one lung cancer, one breast cancer, and three skin cancers). One subject with confirmed skin cancer as recently as 3 years before the interview was unable to recall a date of diagnosis. Of the 25 cases with a date of diagnosis reported by both the medical record and the interviewee, 12 had the same date, and five had a hospital date 1-2 months later than the diagnosis date from the interview. The latter may still indicate no difference in the dates, since hospital dates were discharge dates. The medical record date of five persons was up to 20 months earlier than the diagnosis date from the interview; for three others, the medical record date was 1 to 2 years later.

The overall agreement between responses and records was closer than observed in the analysis of other diseases (except for birth defects). One explanation for this feature may be that people are far more concerned with cancer and birth defects than with other diseases. Given this concern, it is remarkable that the false negative interview responses included one case of lung cancer and one of breast cancer. A false negative interview response rate of 5/34 or 11%, and a (probably) false positive response rate of 4/19 or 21% demonstrate the need for validation of interview responses.

In view of the public's concern, it is noteworthy that three of the four probably false-positive responses referred to leukemia. The laboratory tests of these subjects did not indicate leukemia. Their white blood cell counts were 10,500, 12,100, and 14,000 thousands/mm³ (the upper limit of the range of normal values was 14,000), with normal differentiation, sedimentation rate, and liver tests, and no prescribed cytotoxic drugs. Misinterpretation of an unexplained high (but still normal) blood cell count may have been the cause of the false-positive responses. However, the possibility cannot totally be ruled out that these individuals were treated in the past in a private office, that no paper trail was left, and that the disease was in a phase of complete remission when the blood tests were performed. At any rate, these three cases of leukemia were ineligible for analysis, as TDIAG was prior to 1970, or five years before TIN, or no date was given.

After deleting the benign lesions, 37 confirmed or unconfirmed cases of cancer remained. Thirty four of these were supported by a positive medical record (CA=1 or 2). Table 8.10 shows the IRs for three case definitions, together with the IRs for whites in the Detroit Metropolitan Area (11). The population-based cancer registry of this area is accepted by the International Agency for Research on Cancer (IARC) for quality of data. The registry is the one closest to the study area. To allow comparison of the IR in the current study population with the Detroit IR, the IR for invasive nonskin cancer (to which the Detroit data apply) was directly age-standardized, using the IARC-recommended World Standard Population and method for computing confidence limits for the age-standardized IR (21). The Poisson distribution was used for the confidence limits of the crude IRs (6). The table shows that the IR for invasive nonskin cancer in the study population is lower than in Detroit, but the difference is statistically insignificant. One might argue that this could be explained by some failure to retrieve cases. However, for all other diseases the IR and prevalence were equal to or higher than the rates for the general population, and there is no reason why cancer would be the only exception. It is possible that cancer is so much in the public attention that study participants were less likely to show a recall deficiency than in the case of other diseases.

Table 8.10. Crude and age-standardized (World Standard Population) cancer incidence rate per 100,000/year for the period 1970-1985.

	Male Population		Female Population	
	IR	95% limits	IR	95% limits
All cancers (CA= 1, 2 or 3)	311	174-514	376	233-574
Confirmed invasive cancer (CA= 1)	269	143-460	321	190-507
Confirmed invasive nonskin cancer	186	85-353	231	123-395
Age-standardized IR of confirmed invasive nonskin cancer	173	52-294	221	98-343
Age-standardized IR of confirmed invasive nonskin cancer-Detroit	285		242	

Table 8.11 Odds ratio (OR) and relative risk (RR) for all cancers of *Males* (CA = 1, 2, or 3). TDIAG is corrected for a lag time of 1 or 3 years. A lag time of 3 years decreased the number of eligible cases by one case in the exposed cohort.

MALES: Exposure expression is TAE or having lived in the contaminated AREA (lag time 1 year)

Cancer	Contam. AREA		Exposure = TAEVOC (VOCs)				Exposure = TAEVCL (VOC+Chl)							
	No	Yes	Exposed		Low		High		Exposed		Low		High	
			No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No	220	114	234	100	234	49	234	51	110	224	110	170	110	54
Yes	10	5	9	4	9	2	9	2	7	6	7	4	7	2
OR =	0.97		1.04		1.06		1.02		0.42		0.37		0.58	
p =	0.60		0.58		0.60		0.62		0.11		0.10		0.40	
RR =			1.12		1.20		1.05		0.53		0.48		0.58	
p =			0.53		0.53		0.60		0.19		0.18		0.49	
Lag time 3 yr.			0.78		1.06		0.51		0.35		0.37		0.29	
OR =			0.50		0.60		0.45		0.07		0.10		0.21	
p =			0.54		0.54		0.46		0.13		0.18		0.28	
Lag time 3 yr.			0.84		1.18		0.53		0.44		0.47		0.35	
RR =			0.54		0.54		0.46		0.13		0.18		0.28	
p =			0.54		0.54		0.46		0.13		0.18		0.28	

MALES: Exposure expression is DOSE (lag time 1 year)

Cancer	Exposed		Exposure = DOSEVOC (VOCs)				Exposure = DOSEVCL (VOC+Chl)					
	No	Yes	Low		High		Low		High			
			No	Yes	No	Yes	No	Yes	No	Yes		
No	239	88	239	35	239	53	119	208	119	151	119	53
Yes	8	4	8	2	8	2	7	5	7	5	7	2
OR =	1.36		1.71		1.13		0.41		0.34		0.60	
p =	0.42		0.38		0.53		0.11		0.10		0.41	
RR =	1.51		1.99		1.22		0.53		0.45		0.73	
p =	0.35		0.31		0.53		0.21		0.19		0.52	
Lag time 3 yr.	1.02		1.71		0.56		0.33		0.34		0.30	
OR =	0.61		0.38		0.50		0.07		0.10		0.22	
p =	0.61		0.38		0.50		0.07		0.10		0.22	
Lag time 3 yr.	1.13		1.95		0.61		0.42		0.45		0.37	
RR =	0.54		0.32		0.53		0.13		0.19		0.30	
p =	0.54		0.32		0.53		0.13		0.19		0.30	

Table 8.12 Odds ratio (OR) and relative risk (RR) for all cancers regardless of Females (CA=1, 2, or 3). TDIAG is corrected for a lag time of 1 or 3 years. A lag time of 3 years, decreased the number of eligible cases by one case (exposed cohort).

FEMALES: Exposure expression TAE or having lived in the contaminated AREA (lag time 1 year)														
Cancer	Contam. AREA		Exposure = TAEVOC (VOCs)				Exposure = TAEVCL (VOC+Chl)							
	No	Yes	Low Exposed		High Exposed		Low Exposed		High Exposed					
			No	Yes	No	Yes	No	Yes	No	Yes				
No	257	121	269	109	269	53	269	56	122	256	122	198	122	58
Yes	10	10	10	74	10	52	10	2	4	13	4	11	4	2
OR =	2.12		1.73		2.54		0.96		1.55		1.69		1.05	
p =	0.08		0.20		0.10		0.64		0.32		0.27		0.63	
RR =			1.71		2.74		0.88		1.66		1.87		1.03	
p =			0.20		0.07		0.61		0.27		0.21		0.64	
Lag time 3 yr.														
OR =			1.23		2.03		0.48		0.95		1.11		0.42	
p =			0.45		0.20		0.42		0.57		0.55		0.38	
Lag time 3 yr.														
RR =			1.23		2.25		0.44		1.03		1.24		0.41	
p =			0.45		0.15		0.37		0.60		0.74		0.36	

FEMALES: Exposure expression is DOSE (lag time 1 year)

Cancer			Exposure = DOSEVOC (VOCs)				Exposure = DOSEVCL (VOC+Chl)							
			Low Exposed		High Exposed		Low Exposed		High Exposed					
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes				
No	279	93	279	53	279	40	142	230	142	188	142	42		
Yes	11	6	11	4	11	2	5	12	5	10	5	2		
OR =			1.64		1.91		1.27		1.48		1.51		1.35	
p =			0.25		0.22		0.51		0.32		0.32		0.51	
RR =			1.67		1.94		1.19		1.58		1.64		1.31	
p =			0.22		0.17		0.53		0.27		0.26		0.51	
Lag time 3 yr.														
OR =			1.09		1.44		0.63		0.93		1.01		0.56	
p =			0.54		0.40		0.63		0.54		0.61		0.51	
Lag time 3 yr.														
RR =			1.12		1.62		0.59		0.99		1.11		0.54	
p =			0.52		0.33		0.51		0.59		0.54		0.48	

Of the 37 cases of CA = 1, 2, or 3, 30 were eligible for analysis (13 men, 17 women); that is, TDIAG was later than 1-1-1970 and later than TIN, using one year lag time to correct TDIAG (Chapter 6). If a more realistic lag time of three years was applied, the number of eligible cases was reduced to 27. In either situation, the number of cases was too small for a multivariable analysis separately by sex. The results of a univariable analysis, using ORs and RRs as measures of association, computed as described in Chapter 6, are shown in Tables 8.11 and 8.12 for the broader case definition (CA = 1, 2, 3), and for lag times of 1 and 3 years.

Tables 8.11 and 8.12 show an inconsistent OR and RR pattern. Increased ratios have been found for women only; the ratios for men are equal to or below unity. The only consistent finding is a negative dose-response relationship, if the ORs and RRs for dichotomous exposure are above unity; this is also observed for most ratios below unity. The insignificance of these findings is reflected in large p-values and in the large changes in the ORs and RRs, when just a few cases (one exposed male, two exposed females) became ineligible for analysis with the increase in TDIAG lag time to 3 years. It can be concluded from these results that there is no evidence that VOCs, at the levels which prevailed in drinking water, were associated with an increased risk of developing cancer. Changing the case definitions (including skin cancer, all invasive cancer, and all confirmed cancers), or restricting the population to the Verona exposed and the Calhoun reference cohort did not yield significantly different results.

From the public's local viewpoint, a perceived excess of leukemia mortality in the study population was an issue of particular concern. The three confirmed cases of leukemia (two exposed, one unexposed), are equal to an age-adjusted IR of 49.0 per 100,000/yr (confidence limits 0 - 105). This is higher than expected from the Detroit rate of 10.5, and may suggest an effect of exposure. The difference between the two rates is not significant ($p > 0.2$), however. The observed number of cases is so small that shifting just one case from the exposed to the unexposed cohort would result in an excess of confirmed leukemia among *unexposed* people!

8.14 Skin Diseases and Allergies

VOCs are well known skin irritants, causing skin burns and chronic dermatitis, particularly after prolonged exposure to high concentrations or to undiluted compounds. Although not particularly known as allergens, VOCs may also cause allergies as many other chemicals do. The combined data from the interview and medical records showed that 303 people (40% of the total study population) had at least one skin disorder or allergy, resulting in 406 case-reports. People often listed multiple manifestations of the same disorder as different diseases. In an effort to edit the variety of diagnoses from interviews and medical records, the 406 case reports were grouped into more or less homogeneous categories.

Most of the case-reports (60%) were unconfirmed by a medical record. It is likely that in many cases a skin disorder was not serious enough to consult a physician, or that many physicians may have neglected to enter common disorders, such as diaper rash or acne, in the patient's file. Due to the low percentage of cases confirmed by a medical record, data validation would have rendered the number of valid and eligible cases too small for analysis. Therefore, no validation process took place, and all cases were accepted for analysis if they met the general eligibility criteria of TDIAG \geq 1-1-1970 and TDIAG $>$ TIN. Many interviewee-reported TDIAG were quite likely incorrect; it cannot be expected that the TDIAG of a rather innocent and very common disorder will be recalled correctly farther back than a few years. In general, the earliest date recorded was selected as TDIAG in the analysis. If the interview-derived date was before 1980, while the medical record showed a somewhat later date, the latter was accepted. In the analysis, TDIAG was adjusted for a lag time of 6 months, rather than the one year used for the chronic diseases.

1. Hives : any mention of hives and urticaria.
2. Rash : any mention of a rash except if it was of a viral nature.
3. Acne : any mention of acne.
4. Psoriasis: any mention of psoriasis.
5. Eczema: any mention of eczema or dermatitis (2 cases of seborrhoeic dermatitis or dandruff were deleted; 2 cases of acute allergic dermatitis were grouped under unspecified allergy).
6. Other : any disease not clearly belonging to one of the above categories (neurodermatitis or prurigo, rosacea, dyshyrotic eczema, etc.).
7. Specified : any allergic reaction to drugs or chemicals, insect allergy bites, plants, etc..
8. Asthma and/or : any mentioning of asthma (chronic obstructive hayfever pulmonary disease, as concluded from the presence of emphysema and/or chronic bronchitis, was excluded), hay fever, allergic rhinitis or sinusitis, and allergic conjunctivitis.
9. Unspecified : any other mention of allergy not covered by any allergy of the above categories.

The percentage of cases with missing TDIAG (8.1%) was strikingly high compared with 0.3% for all the analyzed chronic nonskin diseases combined. The analysis of the skin disease category is of dubious value due to 1) the large number of cases rejected for analysis because of a missing TDIAG, and 2) the inability to validate the case diagnosis and TDIAG. Despite this, a univariable analysis was done for completeness, but no RRs, IRs, or prevalence have been calculated; no multivariable analysis was carried out. The results of the analysis are summarized in Table 8.13.

In calculating exposure values, it was recognized that TAE and DOSE were not intended to serve as measures of exposure in the analysis of skin diseases and allergies. The ingested amount of VOCs (DOSE) is probably irrelevant to skin diseases. As for allergies, it is open to question whether the initiation of hypersensitivity and/or the triggering of an allergic attack are dose-dependent. It is unknown whether TAE would correctly reflect the insult to the skin in a dose-responsive manner. However, TAE appeared the best of the possible options and is certainly more responsive than if exposure is defined as "having lived at one time in the contaminated area."

Table 8.13 shows increased ORs for skin rash, psoriasis, and eczema, in relation to TAEVOC, but the p-values are as large as 0.3 or larger. The dose-response effect was negative for eczema. When additional exposure to chloroform (TAEVCL) was taken into account, there was a significant deficit of rashes, leaving psoriasis as the only disorder with an elevated OR, albeit still with a large p-value due to the small number of cases. This inconsistency in the results suggests that the increased ORs for some disorders are a chance effect. This is supported by the significantly decreased ORs for hives and asthma in relation to TAEVCL.

In summary, although statistical analysis has little meaning due to the lack of reference data for validation of the interview responses, a crude analysis shows no evidence of a positive association of VOC-exposure and skin or allergic disorders. The fact that the ORs were

invariably above unity when exposure was defined as "having lived in the contaminated area" illustrates the impropriety of this exposure expression for risk assessment.

Table 8.13 Odds ratios for skin and allergic disorders. The p-values were derived from a one-sided Fisher's exact test.

	Number cases *	OR p	Contam. AREA	Exposure = TAEVOC			Exposure = TAEVCL		
				0-1	Low	High	0-1	Low	High
Hives	37	OR p	1.42 0.19	0.56 0.12	0.49 0.18	0.63 0.28	0.36 0.003	0.36 0.004	0.39 0.06
Rashes	27	OR p	4.21 <0.001	1.46 0.32	1.15 0.49	1.78 0.18	0.39 0.01	0.25 0.003	0.88 0.51
Acne	24	OR p	0.27 0.02	0.10 0.002	0.20 0.06	0 0.01	0.68 0.24	0.89 0.46	0 0.03
Psoriasis	13	OR p	1.74 0.24	1.51 0.33	1.24 0.52	1.77 0.31	1.65 0.33	1.51 0.40	2.11 0.30
Eczema	51	OR p	1.78 0.03	1.24 0.29	1.85 0.07	0.69 0.31	0.93 0.46	1.03 0.53	0.59 0.22
Specified skin allergies	40	OR p	1.20 0.34	0.89 0.45	0.65 0.30	1.12 0.47	0.65 0.13	0.59 0.10	0.85 0.46
Asthma and/or hay fever	49	OR p	1.29 0.54	0.88 0.42	0.84 0.46	0.92 0.52	0.50 0.02	0.47 0.01	0.60 0.17

* TDIAG eligible for analysis, but they likely include *invalid* cases.

8.15 Other Diseases and Conditions

The survey yielded data for a large number of disorders which have not been analyzed for an association with VOC exposure. These are mostly conditions which are unrelated to chemical exposure (heart conditions, arthritis, infections, etc.), or occurred too infrequently (less than 10 case reports prior to validation) to justify the considerable time involved in editing, validating, and analyzing the data. Tables 8.14 and 8.15 list these disorders according to the source of information. These data are provided for the sake of completeness. They do *not* permit an evaluation of the association between disease and exposure, because the cases have not been validated.

Table 8.14. Interview-based disease frequency since 1970 not analyzed in the context of VOC exposure.

Disease	Number of Males		Number of Females	
	Exposed	Control	Exposed	Control
	119	230	132	268
Digestive Tract				
hepatitis, cirrhosis	1	2	2	6
jaundice	9	7	5	5
other	5	3	4	1
Nervous System				
stroke	2	11	2	4
weakness, paralysis	11	11	6	4
emotional problems	10	12	12	21
other	3	2	9	4
Urinary Tract				
urinary tract infection	10	19	42	52
protein in urine	2	2	2	6
other urinary tract	6	5	5	7
Ill-defined Symptoms	65	83	89	153

8.16 The Effect of Tracking on Disease Frequency

Tables 8.16 and 8.17 show the odds ratio (OR) of tracked (former residents of the study area) versus current (residents at the time of interview) people for disease frequencies as measured from medical records and interview responses. The listed frequencies are raw data, not validated or edited other than grouping in broad categories, and they *do not cross-reference* to other data presented for prevalence rates or in OR-tables. Validating the data, for the purpose of illustrating the effect of tracking, would have required much more time and resources than could be justified in a study focusing on the effects of exposure to VOCs. An OR of approximately 1 is expected if tracked and current cohorts have a similar disease prevalence. The nearly equal number of increased and decreased ORs in Table 8.16 suggest that, if the case-finding survey based on medical records were incomplete (and there is no indication it was), incompleteness would have been unrelated to whether or not people moved out of the study area. Table 8.17 depicts a different pattern for case-reports based on the results of the interviews. While there is still an equal number of increased and decreased ORs for the reference cohort, there is a clear tendency towards higher OR values for the exposed cohort. This suggests that tracked people from the exposed group recalled more diseases than did tracked people from the reference group. The differences between former and current residents of the study area depicted in Tables 8.16 and 8.17 are probably real. The number of statistically significant ORs is too large to be attributable solely to the multiple comparison effect. Thus, tracking of former residents appears to be an essential component of epidemiologic studies, although it increases the cost of the study.

Table 8.15 Frequency of disorders since 1970, derived from hospital and physician's records, not analyzed in the context of VOC exposure.

Disorder	Number of Males		Number of Females		ICD categories
	Exposed	Control	Exposed	Control	
	119	230	132	268	
Digestive Tract					
inflammation	2	12	7	22	530 540 555 558 588
other	8	5	7	10	527-8 557 560 562 564 565 569 577 750
Respiratory Tract					
chron. obstr. dis.	4	10	8	7	490 492 493 496
other	3	5	0	2	511-2 515 769
Circulatory Tract					
heart	13	18	13	17	394 410 413-5 424-9 745
arteries, veins	4	9	4	11	346 440-1 446 451 454
Nervous System					
brain and cord	4	6	4	10	331-3 340 342 344 349
peripheral system	2	4	2	4	350 355-7
mental disorders	15	15	12	17	290-319
Genito-urinary Tract					
bladder, urethra	3	3	2	1	597 599
female organs			7	20	614 616-7 620-3 626-8
male organs	3	12			600 604-5 608
Bone and soft Tissues					
spinal joints	11	25	11	20	720-4 756 839
other joints, bone and soft tissues	12	19	10	36	274 524 714-7 733 725-9
abdominal hernia	7	11	4	7	550 552-3
Breast	1	1	3	3	610 611
Ear and Eye	3	8	2	7	361 365-6 372 379 386 389
Neoplasms (benign)	2	5	7	18	210-2 214-9 228 235 239 759
Chest/Abdominal Pain	4	10	9	10	786 789
Ill-Defined Symptoms	8	17	17	30	various

- 1) The numbers of health events in this table refer to numbers of people with the disorder rather than the number of events. One person may have suffered from more than one disorder within the same category.
- 2) Detailed information on the ICD codes is given in the International Classification of Diseases, 9th ed. (DHHS publication PHS-80-1260).

Table 8.16 Disease frequency expressed as the number of cases (N) and the odds ratio (OR) for tracked (former) over current residents. There were 251 (of which 97 were tracked) exposed and 498 (162 tracked) unexposed study participants. The cases in this table refer to any entry in a medical record, and entries have not been corrected or validated.

Disease category	Exposed		Unexposed		Total	
	Cases	Track	Cases	Track	Cases	Track
	N	OR #	N	OR #	N	OR #
Peptic ulcer	7	0.63	17	0.63	24	0.62
Gall bladder	15	1.42	18	0.40	33	0.82
Digestive tract: other	25	0.72	50	1.44	75	1.14
Diabetes mellitus	12	0.51	15	0.14 *	27	0.32 *
Lung	15	0.38 **	24	1.51	39	0.94
Heart	23	1.51	35	0.70	58	1.00
Hypertension	25	0.47 **	53	0.45 *	78	0.45 *
Nervous system (epilepsy, seizure or convulsion)	3	0 .	10	0.89	13	0.56
Nervous system (other)	13	1.38	33	1.38	46	1.36
Thyroid	7	1.20	14	0.34	21	0.58
Female genital tract	8	1.61	26	0.92	34	1.03
Abnormal pregnancy outcome	8	2.74	10	1.39	18	1.92
Congenital defects	5	1.06	10	0.89	15	0.94
Urinary tract	11	0.58	14	1.58	25	1.07
Dermatitis	18	0.30 *	19	0.96	37	0.59
Rashes, hives	20	1.06	15	0.75	35	0.99
Skin: other	12	1.14	21	0.64	33	0.82
Cancer	14	0.41	20	0.35 **	34	0.39 *
Benign tumors	9	0.19 **	23	0.72	32	0.52 **
Disorders of joints	39	0.50 *	90	0.62 *	129	0.57 *
Hernia abdominal wall	11	0.34	18	0.40	29	0.38
Anemia	11	0.58	15	1.85	26	1.19
Allergy	15	0.78	15	0.75	30	0.80
Alcoholism	3	3.22	4	6.32 **	7	4.80
Ill-defined symptoms	25	1.28	47	1.20	72	1.23
Bacterial/viral diseases	4	0.78	246	0.60 *	250	0.67 *
No entries on file	42	1.23	133	1.49	175	1.36 **

* $p < 0.05$ (one-tail Fisher's exact test)

** $p = 0.05$ to 0.1

track OR is the odds ratio of tracked over current people. An $OR > 1$ indicates that the prevalence among tracked people is higher than among current people. The reverse is true for $OR < 1$.

The above categories have been made to match those in Table 8.17 as closely as possible.

Table 8.17 Disease frequency expressed as the number of cases (N) and the odds ratio (OR) for tracked (former) over current residents. There were 251 (of which 97 were tracked) exposed and 498 (162 tracked) unexposed study participants. The cases in this table refer to *interview responses on direct questions*, and have not been corrected or validated.

Disease category	Exposed		Unexposed		Total	
	Cases	Track OR #	Cases	Track OR #	Cases	Track OR #
	N		N		N	
Gall bladder	15	1.42	18	0.40	33	0.82
Liver	21	0.48	26	0.68	47	0.63
Diabetes mellitus	16	0.71	15	0.51	31	0.65
Hypertension	37	0.46 *	76	0.76	113	0.64 *
Nervous system (epilepsy, seizure or convulsion)	3	3.22	10	1.39	13	1.64
Nervous system (other)	36	1.04	30	0.91	66	1.01
Abnormal pregnancy outcome	36	1.04	30	0.91	66	1.01
Congenital defects	12	0.78	22	0.45	34	0.57
Urinary tract	93	1.11	70	1.63 **	163	1.34 **
Dermatitis	21	1.21	32	0.94	53	1.06
Psoriasis, acne	13	4.07 *	32	1.26	45	1.71
Hives	17	2.41 **	26	0.75	43	1.25
Skin: other	54	0.86	46	0.73	100	0.83
Cancer	21	0.78	17	0.29	38	0.66
Anemia	18	1.01	46	2.04 *	64	1.64 *
Allergy	47	1.36	99	0.93	146	1.06
Alcoholism	3	3.22	5	1.39 **	8	1.91
Psychiatric disorders	22	2.03 **	33	1.20	55	1.51 **
Ill-defined symptoms	154	0.90	236	1.30 **	390	1.18

* p < 0.05 (one-tail Fisher's exact test)

** p = 0.05 to 0.1

Track OR is the odds ratio of tracked over current people. An OR > 1 indicates that the prevalence among tracked people is higher than among non-tracked people. The reverse is true for OR < 1.

The above categories have been made to match those in Table 8.16 as closely as possible.

Chapter 9 Discussion

9.1 Introduction

The MDPH had three options for responding to public requests for a study of the health effects of VOC exposure. The first was to evaluate available data on the toxicity of the VOCs, and to extrapolate the known high-dose toxicity to the very low contamination levels of the groundwater. This is a common approach if a quick response is needed. The MDPH and CDC evaluated the situation and concluded that, although an adverse effect of exposure could not be ruled out with certainty, even the theoretical risk was probably too small to be observable. An MDPH evaluation of mortality data for Battle Creek City showed an excess mortality for some diseases, relative to the State mortality rates. However, this was of little relevance to the situation, because the boundaries of the city did not enclose Verona Park, although the neighborhood was directly adjacent.

A second option was a health survey among the residents of the exposed area, or a cross-sectional survey. This option required more resources and time, but was within the range of possibilities for state and local health agencies. However, the scientific value of such a study was dubious, if the question of whether exposure to VOCs had caused observable health effects was to be addressed.

The third option was to conduct an in-depth, comprehensive epidemiologic study utilizing existing knowledge in a wide array of disciplines for optimal results and conclusions. Obviously, this option went well beyond the resources available at the level of state or local health agencies.

At the CDC and the MDPH, it was decided that applying risk assessment principles to an epidemiologic study would elevate it to a state-of-the-science study. This concept was also reflected in the report of the U.S. Department of Health and Human Services on the risk assessment and risk management of toxic substances (19). The Battle Creek Health Study is the first epidemiologic study specifically designed to address all aspects of assessing the human risk of adverse health effects following exposure to toxic chemicals. A full risk assessment includes exposure and health assessments, an analysis of exposure-effect associations, and an evaluation of the uncertainties involved. These components have been addressed in the preceding chapters on a case-by-case basis. The general principles as they apply to epidemiologic studies have been given in Chapter 2. This chapter discusses: what actually was accomplished; what could be expected from a risk assessment perspective; and how this may have affected the conclusions drawn from the analytical results.

9.2 Exposure Assessment

The exposure assessment component of the study was characterized by a number of features, consistent with risk assessment principles. A summary of these features is given below, followed by a discussion of how each related to what could be expected in the ideal situation.

- 1) The basic element was an integrated concentration-time exposure value TAE (total accumulated exposure, incurred during the period people lived in the study area), estimated for each of the VOCs. The estimates were based on a retrospective estimation of $C = f(T)$, which is the concentration C of a chemical in the groundwater as a function of time T as of T_1 . The latter is the date at which VOC contamination of an individual well began at the detection level of 1 ppb. This date was derived by back extrapolation from $C = f(T)$.

- 2) A composite exposure value was calculated as the sum of chemical specific TAEs, each weighted by a measure of dose-equivalence for chronic low-dose toxicity (the NOEL). This composite value was expressed as TAEVOC (sum of the seven VOC specific TAEs) and TAEVCL (sum of TAEVOC and TAE for chloroform in chlorinated city water).
- 3) To express DOSE (amount of VOCs entering the human body), TAEVOC and TAEVCL were multiplied by WATER, the amount of unheated tap water consumed at home, to result in DOSEVOC and DOSEVCL.
- 4) Skin and lung absorption of VOCs, as a secondary route of exposure, were analyzed by using a surrogate measure WASH, derived from the frequency and duration of baths and showers.
- 5) Exposure to other chemicals may cause the same kinds of diseases, and may show a differential distribution among the cohorts. This problem was dealt with by the use of covariates in a multivariable analysis.

Only a few earlier examples have been found of attempts to model past exposure data, taking into account the duration of exposure and the changes of the exposure level over time. The results and the methodology of the Battle Creek Health Study can be compared with the Woburn study (55) and the Santa Clara study (16). These are the only other environmental cohort studies of the effects of VOCs in water.

In the Woburn study, no data were available on current or past exposure levels, and no attempts were made to generate these data. Evidence that people had been exposed consisted of the fact that two city wells were found contaminated in 1979 with chemicals; among these were four of the seven VOCs in our Battle Creek study at levels similar to that in Verona Park. In the Woburn study estimates of an individual's past exposure were made on the basis of an exposure "score", using results from a computer model for the distribution of water from the two contaminated wells. The results from this model, developed by the Massachusetts Department of Environmental Quality and Engineering, were expressed in a "score", a measure of the proportion of the residential water supply believed to have originated from the polluted wells.

This approach had several major deficiencies. The model was not calibrated against the actual situation, although it would have been simple to test tap water samples taken at various points in the peripheral network. In Battle Creek, a city similar in size to Woburn, monitoring of tap water was initiated immediately after the detection of groundwater contamination. At the Battle Creek consumer's end of the water network, VOCs were present at levels of at most 8 ppb due to the dilution effect of uncontaminated city wells. The levels further decreased when the most seriously contaminated wells were taken out of operation. The VOC levels in tap water did not show a particular distribution pattern. This contradicts the Woburn claim that a computerized model for VOC distribution through the city mains can predict VOC levels in individual dwellings. Further, such a model is not compatible with the thorough mixing of contaminated water with clean water from the many other wells during storage and chlorination, or other forms of water management.

The second deficiency in the exposure assessment is that the Woburn model yielded estimates for entire blocks of dwellings rather than individuals, and that the amount of water *available* rather than the dose was used as a measure of individual exposure. This required the (unrecognized) assumption that the daily consumption of unheated tap water was the same for each study participant, an assumption which is certainly invalid. As shown in the Battle Creek study, many people may not consume cold tap water at all, rendering them unexposed regardless of the VOC concentrations. The Woburn analysts did not present evidence for the implied assumption that all chemicals present in a one-time sample were also present at constant levels

during the entire assumed period of exposure. It was assumed that T_1 was the same as the date the wells came into operation. This approach ignores basic toxicology and risk assessment principles, rendering the results of the Woburn study difficult to interpret, and probably invalid.

In Santa Clara County, California, an industrial spill of TCA (among other compounds) led to groundwater contamination. The California Department of Health Services conducted a short-term prospective cohort study of adverse pregnancy outcomes (16). Individual exposure levels were estimated using a distribution model for piped water similar to the model used in the Woburn study. Some efforts were made to estimate T_1 , to ensure that the exposed women were exposed during the entire pregnancy period. An attempt was made to validate the results of the water distribution model by comparing model-predicted values with values observed during a field test with fluoride injected in the system. The contractor for this part of the study concluded that the comparison showed "considerable error"; observed zero values were paired to predicted high positive values, and the reverse (63). It is unexplained why the contractor averaged the availability of TCA over 24 hours, arriving at a final exposure value which included high exposure levels at night, a period of minimal or no water consumption, however. During the daytime, when other (clean) wells were pumped and when individual use of tap water peaked VOC concentrations were close to zero. No information was provided on whether and how these errors were corrected, and it may be inferred from the epidemiological study report (16) that no corrections were made. There was also no information as to whether and how TCA concentrations were used in the calculation of individual exposure values. Finally, due to the design of the water distribution model, the exposure estimates were made for blocks of 100 dwellings (defined by a pipe "node"), and not for individuals.

The Santa Clara County study differed positively from the Woburn study: efforts were made to estimate T_1 , to prospectively monitor $C = f(T)$, and to calibrate the water distribution model. As in the Woburn study, however, the analysis was performed with the "available" exposure rather than the dose, and concurrent exposure to other chemicals and exposure through other routes was ignored. The shortcomings in the exposure assessment may have led to the peculiar conclusion that adverse pregnancy outcomes were associated with the amount of municipal tap water used, regardless of the contamination with TCA. Thus, no toxicant was identified.

These two studies are not the only ones in which quantification of individual past exposure data has been attempted. As an essential part of a case-control study of colorectal cancer and chloroform in drinking water, Lawrence and Taylor (56) modeled the expected historical chloroform concentration as a dependent variable in a linear regression on the amount of chlorine used, the effluent chlorine residual, and the type of source water. Data on these regressors were derived from city records. The method is appealing for its relative simplicity and general applicability, and because it allows the use of time in the assessment of individual exposures, although the authors did not utilize this potential of their model. Estimating historical exposure data is a problem not limited to water contamination. In a case-control study of breast cancer and exposure to radiation from repeated diagnostic fluoroscopy for tuberculosis, Boyce and coworkers (10) attempted a quantification of individual radiation doses from decades-old records with data on the kind of device and the procedures used. This yielded a total accumulated radiation (comparable to TAE), enabling a dose-response evaluation. Their results were confirmed by the findings of other studies (43, 45), which showed the credibility of their modeling based on past exposure.

The methods developed for the exposure assessment in the Battle Creek Health Study have several important advantages over the methods used in the studies cited above. These are:

- * T_1 can be estimated for individual wells, which is essential to ensure a proper exposure-disease sequence.
- * Exposure can be expressed in various forms of one-time or accumulated exposure.

- The time-dependency of the exposure is accounted for.
- Exposure can be converted to dose.
- Exposure estimates are individual estimates.
- The exposure assessment enables a dose-response analysis and an evaluation of the effect of various exposure expressions on the analytical results.

Another important feature of the exposure assessment methodology is that it is probably applicable to experimental, environmental, or occupational exposure scenarios in general. There are several weaknesses in the various components of the exposure assessment, however, and application of the method to other scenarios could require some adjustments. However, most if not all weaknesses are due to deficiencies in the data sources and not in the methodology itself. The most important deficiencies come from a lack of standardization in the designs of the animal or human studies which provide the background data, or from a total lack of data. The methodology is designed to accept easy adjustment if new data becomes available, e.g., standardized absorption rates would enable the estimation of a composite dose estimate for all routes of entry combined, and standardized NOEL studies would improve the accuracy of the TAEVOC or TAEVCL estimates.

The time dependency of exposure, $C = f(T)$, was developed from time series of VOC concentrations in well water, using data on city well properties and the results of monitoring city wells for VOCs. The extrapolation of the results to neighboring residential wells was based on the assumption that *once contaminated*, a particular private well will show the same $C = f(T)$ estimated for city wells. Arguments supporting this assumption are provided in Appendix C. This assumption was *not* supported for wells free of the chemical of interest at the time of sampling in 1981-1983, or for the private wells in Dowagiac and Springfield. Factors which are likely to have impaired the accuracy of the estimated $C = f(T)$ include: some arbitrariness in the selection of wells providing raw data for estimating $C = f(T)$, the changes in groundwater management during the monitoring period, the rather limited period of monitoring (up to 2.5 years), and the inability to entirely rule out that additional sources contributed to the aquifer contamination. Calibration of the model using data from a few private wells sampled twice showed a satisfactory degree of accuracy, however (Appendix C). The error in T_1 estimates (several months) was much smaller than potential errors in dates of disease diagnosis (several years). Further, as inaccuracies in $C = f(T)$ affect individual TAEs randomly, and as the estimation of $C = f(T)$ is entirely independent of the health status, inaccuracies will not change the direction of a statistical association between exposure and disease, although that may affect the precision of the estimates of exposure and risk.

The accuracy of the individual exposure estimates of residents in the Dowagiac and Springfield areas is a different issue. The dimensions of the study were not large enough to permit the generation of hydrological data for modeling $C = f(T)$ in these areas. A conventional approach to exposure assessment is to assume constancy over time of VOC levels in the one sample taken from the residential wells. Any other approach would be as arbitrary, given the lack of data. The assumptions that the contamination in these areas started before 1970, and that the VOC concentrations were at an equilibrium since 1-1-1970, are difficult to validate due to the lack of monitoring data. In the case of Dowagiac, solvents had been stored in tanks for decades. The contamination levels were much higher than in Verona Park, despite approximately the same distance between the wells and the contamination source. It was, therefore, reasonable to expect that the contamination in Dowagiac started many years before 1977 (the earliest T_1 in Verona Park), and that this long period was sufficient for the concentrations to reach an equilibrium. Even if the VOC levels were not constant, the results of the analyses, with regard to the direction of exposure-disease associations and the dose-response effect, would not have been substantially different, however. The VOC levels in Dowagiac were much higher than in the other areas and most members of this cohort regardless of $C = f(T)$ would still remain in the highest exposure category.

For Springfield, there were no substantive arguments favoring or rejecting the assumption of constancy and a T_1 before January 1970. However, the number of participants from that neighborhood was too small, and the contamination level too low for this to have a measurable effect on the outcome of the analysis. The weakness of the data for Dowagiac and Springfield was recognized, and they were excluded in alternative analyses. In no case did the result of the alternative analyses change the results of the initial analysis of the total population.

As explained in Chapter 4, lack of physicochemical data made it impossible to estimate the absolute amount of VOCs entering the body along secondary routes. Most important were the lack of data on partition coefficients for water-air, air-blood, and water-skin lipids, and the absorption rates for lungs and skin at the very low air concentrations prevailing in bathrooms. Some investigators have advanced the idea that lung and skin absorption rates at bathroom conditions of exposure are similar to those observed in experiments with VOC levels several orders of magnitude higher, assuming constant physicochemical properties with increasing VOC levels (3, 4, 13, 61). This assumption is not supported by the facts, and arguments against physicochemical constancy have been provided in Chapters 3 and 4. Since constancy of partition coefficients or absorption rates has been studied only for a very narrow range of high concentrations, extrapolating these results many orders of magnitude below this range seems questionable.

In this study, the problem of lack of data on physicochemical parameters was circumvented statistically, by entering the secondary exposure variable WASH (Equation 7, Chapter 4) as a covariate in a multivariable model, with or without linkage to the level of VOC contamination. The use of a composite dose estimate for the three routes of entry would have required the indefensible assumption that the amount of VOCs absorbed through the lungs is equally as toxic as the same amount absorbed through the skin or intestines. The advantage of the chosen approach is that it does not require this assumption.

The estimation of a composite oral exposure value for the seven VOCs found in the drinking water was based on the assumption of equitoxicity, after careful consideration of available data on the quantitative aspects of toxicity (Chapter 3). Equitoxicity could not be assumed in concurrent exposure to other chemicals; hence, such exposure was dealt with as a covariate in a multivariable model. Differentiation between chemicals with regard to equitoxicity among chemicals in a situation of multiple exposures was not found in the two environmental studies cited above (55, 56). However, no environmental study is complete unless these issues have been addressed. Individual differences in multimedia and multichemical exposures are so large and unpredictable, that to ignore them may cause large differential errors in the total exposure estimate of individuals. This can not only affect dose-response relationships, but may also lead to biased results. The approach in the Battle Creek study minimized the potential of bias in the exposure assessment and hence in the conclusions.

Quality control and assurance procedures (QA/QC) in risk assessment require validation of the exposure estimates. How this applied to skin and lung absorption rates was discussed earlier. QA/QC for water quality tests confirmed the EPA standards. A logical extension of the QA/QC requirements would be to confirm that currently clean residential wells were also free of VOCs in the past. However, since there were no historic data on these clean wells, it was assumed that currently unexposed people were also unexposed in the past. There are no methods for validating this assumption. The Battle Creek City cohort differed in this aspect from the other cohorts. Water quality data have been available since 1981 because of a monitoring program which also included testing of water samples at the consumer's end of the city's water supply network. The water quality information also contained data on trihalomethanes, almost exclusively chloroform. Since the toxicity profile of chloroform is similar to that of the VOCs proper, it was possible to compute alternative exposure expressions: exposure to the seven VOCs alone versus VOCs plus chloroform (abbreviated as TAEVOC versus TAEVCL, and DOSEVOC

versus DOSEVCL). The analysis of diseases in association with these alternative expressions of exposure is proposed as the approach most consistent with an unprejudiced analysis of the effects of exposure. If a choice is to be made, an analysis using DOSEVCL seems the one most in line with toxicologic facts.

9.3 Health Assessment

A major advantage of case-control studies is the good quality of the health data, at least for incident cases. In a retrospective cohort study, both the quality and completeness of health data are crucial issues. QA/QC procedures must be tailored to data generated in the past. No QA/QC of health information was applied in the environmental studies reviewed in this report, with the exception of leukemia in the Woburn study. In the Battle Creek Health Study, the design had a built-in QA/QC of health data by gathering information from multiple sources on the present and past health status of individuals, and by repeating certain interview questions in different formats. This yielded favorable results: more diseases were identified; the higher quality of the diagnosis allowed the avoidance of duplication for the same disease; and the dates of diagnosis were more reliable.

There were deficiencies in the health assessment, however. The most important one was the omission in the study design of QA/QC procedures specifically tailored to the retrieval and abstraction of medical records. The design relied on the proven expertise of professional, trained registrars. In hindsight, the control in this area was not consistent with the stringent QA/QC procedures used for laboratory tests. QA/QC programs for medical records have yet to be developed. The main problems will be the cost and the legal aspects of obtaining, keeping, and filing more than a very short summary of a patient's record.

The problem of discrepancy between data from interviews and from medical records is not new. Belloc (8) tested the recall capacity of people with respect to hospitalizations during the one year preceding the interview. Of the positive responses, 11% were not matched by a hospital record, and 14% of the positive medical records were not recalled in the interview. Five percent of the interviewees mentioned a recent admission but the record indicated hospitalization more than one year before the interview. Trussel and coworkers conducted perhaps the most comprehensive and largest study ever to address the problem of recall deficiency (93). The study involved an entire community and 13,113 interviews. In addition, 329 physicians reported on 1569 patients. A random sample of 846 participants had a full medical examination with unlimited access to medical specialties, laboratory, and X-ray services. Existing medical records were added to the database. With regard to diseases relevant to the Battle Creek study, Trussel *et al* showed that 15% of the respondents reporting diabetes, and 25% of those reporting malignant or benign growths had no medical confirmation for these conditions. Of the medical diagnoses, 36% of diabetes and 90% of neoplasms were not recalled in the interview. Trussel *et al* concluded that, at best, household interviews about chronic diseases may give a minimum estimate of morbidity. In a second paper, they concluded that discrepancies between the interview and medical data were much more "a reflection of the state of medical diagnosis and patient-physician communication in routine general medical care in the country today" than a recall deficiency of the interviewee (23).

More recently, Gordis (40) postulated that questionnaires rather than responses are deficient, and that standardization was urgently needed. This conclusion is acceptable but it does not eliminate the most important condition for a good epidemiologic study, that both health and exposure data be validated. (The full questionnaire used in this study is attached to this report in Appendix A.)

Although less thorough than the study done by Trussel and coworkers, the multiple source approach in the current study was successful in finding cases. The prevalence of most diseases analyzed was substantially higher than reported for the general U.S. population by the NCHS. This may suggest that incompleteness of case-finding in the Battle Creek study was not a problem.

The accuracy of the date of first diagnosis (TDIAG) requires special attention in future studies. Since a proper temporal sequence (exposure precedes the disease) is a major requirement in the epidemiologic analysis, estimation of T_1 and TDIAG is of primary importance. Except for Belloc's study of recall errors (8), no useful reference was found on the issue of the reliability of TDIAG. In the current study, the discrepancies between interview data and medical record data regarding TDIAG were usually on the order of years, and were in some cases more than a decade. There was no clear relation between the magnitude of the discrepancy and an individual's exposure status. The dates reported by the interviewee and the dates from the medical record were closest for cancer and reproductive outcomes. The question of the date of disease initiation rather than TDIAG will be discussed in the next section.

9.4 Statistical Analysis

A proper epidemiologic analysis entails more than just a statistical analysis. In general, too little attention has been given to criteria of causality in the interpretation of health studies of chemical exposure. Formulated and updated by Hill (44), these criteria are oriented more toward biology than statistics. The most important criteria are: a) proper temporal sequence of exposure and effect, b) positive dose-response relationship, and c) biological plausibility of the effect. The development of analytical models in this study, and the evaluation of their results, were definitely oriented towards Hill's criteria, and the time factor has become the most crucial component of the statistical analysis.

With a point source of contamination, it is inevitable that there must be a T_1 , and the level of concentration must increase over time until equilibrium is reached. As long as the spill continues, no declining concentration levels are expected. Hence, there must have been an initial period in which people were unexposed because they lived in the area before T_1 , and concentrations must have increased over time to reach the sample levels. It was also shown that T_1 and $C = f(T)$ are related to specific chemical compounds and that T_1 also is well specific. It is because of this specificity that the individual composite exposure to the seven VOCs and chloroform could not be expressed as a simple function of time of residence or the sample concentration of a single chemical. Thus, modification of PHGLM, a statistical package used for the proportional hazard analysis (41), became an essential prerequisite for a proper exposure-effect analysis.

If a disease occurs frequently enough to allow a multivariable analysis, the proportional hazard model is the natural tool for retrospective cohort studies (50, 83). This model considers the risk in association with TFU, the duration of followup. Its modification to meet the specific needs of exposure levels changing over time has not been described earlier. Various important features render this modified model generally applicable. It allows the incorporation of any form of $C = f(T)$ for any number of chemicals, without the condition of uniformity among these chemicals. The modification permits the inclusion of lag time in TDIAG, even if it varies over time and it also offers a choice between calendar time or individual time in computing TFU.

The latency period before a disease becomes clinically detectable is another important time-related issue in evaluating the temporal sequence of exposure and effect. This period involves two different components. The first is the minimum period of exposure required to induce an irreversible first stage of the disease (exposure-time threshold). The second component

is the latency period proper, the time lapse from the initiation of the disease until it is clinically recognizable. It would be against sound scientific reasoning to assume that the initiation of cancer, hypertension, etc., begins the same day as exposure. Some time must elapse for pathogenetic processes to evolve and become complete. Yet, the zero-threshold principle, in which any finite exposure (regardless of duration and level) causes a finite risk, has become government policy in the risk assessment of cancer (78). Thresholds above zero for other health effects have been accepted almost universally, but this has so far been formalized only in the values for ADI (Allowable Daily Intake). ADI values stipulate a threshold for the amount of exposure, but not for the duration of exposure.

Since both time and dose threshold are intrinsic components of the computer algorithm used to estimate exposure values and the temporal sequence for exposure and disease, the analytical model used in this study accepts any threshold values. Changes can easily be made if new toxicologic data become available. The exposure-time threshold in the current analysis was set at zero months because little, if anything, is known about the minimum threshold and how this may vary among diseases and chemicals. A lag time of 1 year was used for all diseases except pregnancy outcomes and skin disorders. For cancer, a 3 year lag time was used in an alternative analysis. This lag time is an assumed value for the sum of the latency period proper, the time delay in diagnosis once a disease became observable, and the potential error in recorded or reported TDIAG. For pregnancy outcomes, the lag time for correcting TDIAG cannot be longer than the time since conception. The nature of skin diseases is more subacute than chronic; hence, a lag time for TDIAG of 6 months was used. The use of a lag time is recognition that some degree of error has to be taken into account, and a value of 1 year is hardly more than symbolic, as the actual lag time is unknown for any chronic disease.

The proportional hazard model cannot be used for the analysis of reproductive events, as *the period-at-risk does not end with the pregnancy*. The logistic regression model is the logical choice, provided that the unit of observation is the pregnancy and not the woman. This is not only because risk increases with the number of pregnancies per woman; it is also because the exposure status for each pregnancy can be defined. The exposure status can be defined as exposure accumulated between the index pregnancy and the preceding pregnancy, or as exposure accumulated from the very start of the exposure. The first option is biologically more relevant, since the fetus is more likely to be affected by exposure during the pregnancy than by earlier maternal exposure. However, since it cannot be precluded that maternal toxicity during a short period preceding the pregnancy may affect the fetus, the exposure accumulated during the pregnancy plus the interpregnancy accumulation (the exposure value used in the current study) seems to be a reasonable compromise.

There are serious controversies about the meaning of a p-value in a study with multiple comparisons. Breslow and Day (12) have expressed what is probably the prevalent opinion: in the absence of an *a priori* hypothesis about the effect of exposure, conventional p-values are unreliable indicators of the likelihood that chance has caused the observed effect. For each 100 diseases analyzed, five are likely to show a statistically significant association with the exposure at $p=0.05$. With regard to the current study, multiple comparisons were made not only because of the many disorders analyzed, but also because multiple exposure expressions were used.

Several corrections have been proposed to adjust the p-value for multiple comparisons, but none have gained acceptance in the evaluation of statistical outcome. For instance, an often proposed correction for alpha (the significance threshold p-value) is $\alpha = 1 - (1 - A)^n$, in which α is the proposed new alpha, A the conventional alpha, and n the number of comparisons made. This would in most cases not be justified, since the associations tested have usually not been selected at random, but are based upon certain biological considerations. In this study, much less weight was given to the p-value than to the consistency of the measures of associations (OR and RR) and the dose response relationship across statistical models, strata, and exposure expressions.

One of Hill's major criteria of causality is a consistently positive dose-response relationship (44). It is the only applicable criterion in studies which are too small to render OR or RR statistically significant. The Battle Creek Study population of 749 (all ages) is small indeed for a study of chronic diseases, although this study size is quite large for an analytical environmental study.

The dose-response relationship preferably is studied in a regression model. For most diseases in the Battle Creek Study, however, the incidence was too low to permit a multivariable analysis. Fourfold tables (stratified by exposure level) were used instead. Such tables have the serious limitation that they offer little or no control over confounding, and the duration of observation is ignored if the OR is used as a measure of effect. The correlation between ORs and RRs was found to be very close with regard to both the magnitude and the direction of the dose-response relationship, however. Stratification does not offer a solution for confounding, because the number of observations in the "case cells" very quickly approaches zero with each stratum added. Thus, except for hypertension, not a single disease could be analyzed with stratification for more than the two levels of exposure and two levels of the single most important potential confounder, e.g., gender or age. On the other hand, to refrain from any analysis because of this limited number of strata would certainly be a less defensible alternative from a public health standpoint. For diseases with too few cases to allow stratification of exposure levels, no stratified analysis was attempted a dichotomous exposure value was used.

In using a multivariable model, it must be kept in mind that most covariates reflect a current status, while the model assumes the status in the past. Certainly, sex, education level, and (in most cases) occupation, would not change in time if the followup period starts at an adult age; the same is probably true for the current status of smoking and alcohol use expressed as dichotomous variables. However, it is not possible to extrapolate into the past from current test results to historical liver and kidney function tests, serum values of exogenous chemicals, etc. Hence, in the search for risk factors other than exposure to VOCs, care was taken to select covariates of which current values can be assumed either to change little over time, or to be "predictive" of past levels. Some examples are the body mass index (QI), income, and serum levels of triglycerides or DDT, especially if the period of time to be covered is relatively short (10 years or less). This was true for the vast majority of cases. Yet, caution remains necessary in making inferences from the results of multivariable analysis. Therefore, in the final interpretation, much weight has been given to consistency between the results of univariable and multivariable analyses.

Biological plausibility is a criterion difficult to apply in the analysis of the study results. It would be an error to conclude that a disease cannot be associated with exposure simply because nothing is known about that association. Most chronic diseases have not been tested in the context of chemical toxicity, and the mere lack of biological plausibility should not be used as the only argument for rejecting a statistical association. This criterion was used more liberally when setting priorities for selecting health effects to be analyzed. For instance, although there seem to be a large enough number of cases with arthritis, the disease was not selected for analysis because it was judged not to be a biologically plausible effect of exposure to VOCs.

Consistency of the findings, another important criterion formulated by Hill (44), can be evaluated by comparing the study outcomes for different strata. This criterion was also applied to cover consistency across populations, analytical models, and exposure expressions. Although the many alternative analyses may seem to be "overkill", they provide an opportunity for evaluation of exposure-disease associations utilizing multiple, more-or-less independently obtained, results. If only one analytical model and exposure expression were selected, the analysis could be biased by personal prejudice, especially in sensitive situations or controversial issues. The use of multiple analytical models avoids this problem.

9.5 Results

A summary of the frequencies of diseases which have been analyzed is given in Table 9.1. The number of cases refer to all cases encountered, regardless of the date of diagnosis, and whether or not the diseases were detected while the individual was living in the study area. Thus, these numbers are what would have been the result if this study were a simple frequency survey. They are probably close to what the public perceives as a disease frequency. Whether or not differences in the rates between the populations are statistically significant is indicated by a low (lower than 0.1) p-value for the *ratio* of the frequency rates. The number of disorders with higher rates in the exposed area is about the same as the number with lower rates, as one could expect from cohorts formed on the basis of a factor (VOC exposure) unrelated to disease occurrence. However, if the comparison is limited to rate ratios that are statistically significant, five out of six disorders are more frequent among residents of the contaminated area.

Table 9.1 Overview of disease frequencies.

Disease		Diagnostic Category	Contaminated Area		Reference Area		p-value
			Cases	rate/1000	Cases	Rate/1000	
Diabetes	*	1-4	13	51.8	15	30.1	0.10
Hypertension	*	1-2	33	131.5	82	164.7	
Gallbladder	*	1	15	59.8	19	38.2	
Hypothyroidism	*		6	23.9	14	28.1	
Epilepsy	*		4	15.9	10	20.1	
Kidney disease	*		6	23.9	12	24.1	
Peptic ulcer	*	1	7	27.9	15	30.1	
All cancers in men	*	1-3	5	42.0	10	43.5	
All cancers in women	*	1-3	11	83.3	11	41.0	0.07
Birth defects	**		8	33.5	19	39.5	
Miscarriage	**		46	161.4	42	80.3	<0.001
Prematurity	**		8	33.5	21	43.7	
Stillbirth	**		5	20.9	8	16.6	
Low birth weight	**		18	76.3	41	85.8	
Skin - hives			23	91.6	32	64.3	
- rashes			21	83.7	14	28.1	0.001
- acne			5	19.9	27	54.2	0.02
- psoriasis			11	43.8	11	22.1	0.08
- eczema			31	123.5	42	84.3	
Specific skin allergies			21	83.7	48	96.4	
Asthma or hay fever			31	123.5	53	106.4	

The number of cases is *unrestricted*, no eligibility criteria were applied. Therefore, the numbers do *not* compare with those in OR and RR tables, and the rates/1000 population are *not* incidence or prevalence rates! The p-values were derived from a one-sided Fisher's exact test for the *ratio* of the rates.

- * There was sufficient information for case validation; the diagnostic categories (second column) are explained in the Appendices and, for cancer and peptic ulcer, in Chapter 8. For other disorders, case validation was limited or (for skin diseases and allergies) not possible at all.

** The denominator for miscarriage is the number of all pregnancies; the denominator for birth defect/prematurity/stillbirth/low birth weight is the number of pregnancies *not* ending in miscarriage. For birth weight, the denominator is somewhat smaller due to missing values.

This is more than might be expected if chance were the only factor determining differences in rates. This raises the question of whether exposure to VOCs could have been one of the possible causes. To answer this question, we must go beyond the simple frequency survey.

The first necessary correction is to *delete cases diagnosed before they moved to the study area*, since these cases could not possibly be attributed to the VOC contamination. The second necessary correction is to *accept cases only if diagnosed in 1970 or later*. The reason for this criterion of eligibility is that the search for medical records was limited to the period 1970-1985; a search for records covering the entire lifespan of residents (which may go back as far as 1895) was obviously not feasible. Thus, information on diseases diagnosed prior to 1970 was derived totally from interview responses. This source of data has a high likelihood of being inaccurate, or biased towards better disease recall by people in the exposed area. If both corrections are applied, a changed pattern of differences in disease rates emerges, as shown in Table 9.2.

Table 9.2 Overview of odds ratios (OR) and rate ratios (RR) for dichotomous exposure expressions. An asterisk indicates statistical significance at $p < 0.1$. For details, see Chapter 8 and the appendices relating to specific diseases.

Disease and number of cases eligible for analysis ()	Exposure = VOCs only				Exposure = VOCs + CHL			
	TAEVOC OR	DOSEVOC RR	DOSEVOC OR	TAEVOC RR	TAEVCL OR	DOSEVCL RR	DOSEVCL OR	TAEVCL RR
Diabetes (20)	1.99*	2.06*	2.60*	2.77*	0.73	0.86	1.04	1.21
Hypertension (92)	0.49*	0.54*	0.49*	0.56*	0.55*	0.66*	0.59*	0.70*
Gallbladder disease (28)	0.85	0.89	0.91	0.97	0.34*	0.41*	0.41*	0.48*
Cancer men (13)	1.04	1.12	1.36	1.51	0.42	0.53	0.41	0.53
Cancer women (17)	1.73	1.71	1.64	1.67	1.55	1.66	1.48	1.58
Epilepsy (6)	0.47	0.49	0.56	0.61	2.44	2.83	3.01	3.52
Hypothyroidism (12)	1.18	1.24	1.05	1.18	0.48	0.57	0.49	0.48
Kidney stones (7)	0*	0*	0	0	0.65	0.75	0.80	0.94
Peptic ulcer (23)	0.83	0.88	0.82	0.90	0.76	0.89	0.87	1.02
Low birth weight (59)	0.45	0.55	0.72	0.81				
Miscarriage (88)	0.98	0.87	0.82	0.77				
Prematurity (7)	0	0	1.23	1.39				
Birth defects (11)	0.45	0	1.61	1.50				
Hives (37)	0.56	0.67	0.36*	0.45*				
Skin rash (27)	1.46	1.23	0.39*	0.35*				
Acne (24)	0.10*	0.12*	0.68	0.70				
Psoriasis (13)	1.51	2.37	1.65	6.10*				
Eczema (51)	1.24	1.29	0.93	0.97				
Specified allergies (40)	0.89	0.82	0.65	0.62				
Asthma or hay fever (49)	0.88	0.93	0.50	0.51				

This study has not identified any adverse effect of exposure to VOCs through contaminated drinking water. *This conclusion was not reached lightly.* Rather than relying on a single

expression of exposure or analytical model, the data were evaluated from several different angles. ORs and RRs consistently lower than unity indicate that the conclusion of no-effect could *not* have resulted from an insufficiently large study size. As shown in Table 9.2, ORs and RRs are less than one for hypertension, gall bladder disease, kidney disease, peptic ulcer, miscarriage, stillbirth, low birth weight, allergies, acne, and hives. Positive associations of VOCs with epilepsy, hypothyroidism, congenital defects, cancer, skin rashes, and eczema were rejected because of 1) a consistently negative dose-response relationship and 2) elevated ORs and RRs for some exposure expressions, and decreased values for others.

The results of the analysis of diabetes and premature births were mixed, that is, increased and decreased ORs and RRs with positive or negative dose-response effect. With regard to diabetes, if ORs and RRs were significantly increased, the dose-response relationship was negative. This inconsistency in results led to the conclusion of no-effect. A much larger study would be required to resolve such inconsistencies.

Psoriasis was the only disorder with a positive dose-response relationship and consistently elevated ORs. The number of cases was very small, however (N=13). Based on the preset rules, it could have been concluded that this association indicated an adverse effect of VOCs which did not reach statistical significance because of the low statistical power of the study. Such a conclusion was not reached because skin diseases constituted the only category in which unconfirmed cases were included for analysis. In the majority of skin disease cases, no validation of the diagnosis was possible due to lack of medical records. An analysis of confirmed cases only would not have been meaningful (there were only three validated, eligible cases of psoriasis). The possible positive association for psoriasis was the only such association out of the 19 disorders analyzed. It may have been a chance effect resulting from making multiple comparisons.

A small study size may not only hide a positive exposure-disease association, but may (by sheer chance) generate an OR/RR lower than unity, even with a consistently negative dose-response relationship. For instance, hypertension and acne showed statistically significant ORs and RRs below unity and the dose-response relationship was negative. Based on the preset rules of evaluation, this could have been interpreted as a protective effect of VOCs because of its consistency and the significant p-values. A protective effect may seem illogical, but that does not rule out biological plausibility. The relationship (if any) between a VOC protective effect and these disorders has not been studied in animal experiments, nor in human studies. It was concluded that exposure to VOCs had no protective effect and priority was given to consideration of a chance effect resulting from making multiple comparisons. Further studies are needed to confirm the negative associations, or to identify unrecognized confounding factors responsible for these associations.

The evaluation of associations between the results of the clinical examination and VOCs is complicated by the fact that these results reflect a current status with an unknown relation to time. Abnormal test results could have existed for an unknown period of time, ranging from very recent to decades ago. Thus, it is not possible to identify the exposure-effect sequence. Further, many test outcomes may be influenced by existing diseases and pharmaco-therapeutic drugs. Therefore, the analysis was limited to estimating the ORs for abnormal test values and exposure (defined as having lived in the contaminated area), and to simple correlations between test results and TAE or DOSE as continuous variables. Neither analysis is proper, since they do not validate the time sequence and do not control for the presence of diseases and drug treatment.

With this caveat in place, it may be concluded that there is no clear evidence of an effect of VOC exposure on clinical chemistry outcomes. Of the 18 tests, only one (LDH) had a statistically significant increased OR for Degree I values (less than 10% above the upper normal

limit); this was contradicted by a decreased OR for Degree II values (more than 10% above the upper normal limit). In contrast, five tests showed statistically significant ORs lower than unity: iron binding capacity, hemoglobin, total bilirubin, triglycerides, SGPT, and GGTP (Table 8.2). This is more than one might expect if chance is the only determinant of the direction of the OR. It is inappropriate to analyze current biochemistry data without knowledge of past biochemistry data. Therefore, the only permissible conclusion is that there is no evidence of a (negative or positive) effect of VOCs.

There is abundant evidence of VOC effects in humans, but the exposure levels involved are many orders of magnitude higher than environmental levels. Only two analytical studies have found an adverse effect of ambient environmental VOC exposure. The first, the Woburn study (55), reported that exposed children had an excess of leukemia, perinatal deaths, congenital anomalies, and some other disorders. Reviewers of that study have identified many serious problems in the conduct and analysis of the study (58), but the main deficiencies from the viewpoint of risk assessment have not been mentioned. These include the failure to define a specific exposure, to express exposure in terms of dose, to address exposure to multiple chemicals with different toxicologic profiles, and to consider the lack of consistency in the analytical results. Further, there was no quality control of the exposure assessment. As noted above, it is very likely that the assessment of the individual "scores" of exposure was invalid due to the use of an improper water distribution model. The investigators ignored the fact that an analysis based on improper exposure data, in a situation with a known cluster of disease and the presence of contaminated wells, was bound to yield a positive statistical association. In the Battle Creek study, a similar effect was observed *if exposure was defined as having lived in the study area*. In this context, it is worthwhile to underscore that none of the positive findings in the Woburn study were found in the Battle Creek study (all ORs were below unity) or in the Santa Clara County Study (16).

The second study reporting a positive effect of contaminated well water was that of Clark and coworkers (18), dealing with Hardeman County, Tennessee. They tested the blood serum of 31 exposed people at two points in time, two months apart. At the time of the first test, the subjects were still using the contaminated water; this exposure stopped shortly thereafter. The presence of 13 chemicals was reported, but no attempt was made to assess individual exposures with regard to the specific compounds contaminating individual wells, or the composite effect of multiple chemicals. Among the chemicals were three chlorinated methanes and one of the VOCs (PCE) found in the Battle Creek study. The geometric means of alkaline phosphatase, SGOT, and SGPT were higher at the first testing (p-values 0.04 to 0.0003), although all were well within the range of normal values. Enlarged livers were found in seven people. The study ignored lower values for total albumin, bilirubin, and GGTP (p-value 0.02 to 0.0001)! Further, a difference in means is an improper measure of effect, as it does not account for the relationship between the test outcome and each individual's exposure status. This is a typical example of a situation where application of Hill's causality criteria (44) of internal consistency and biological plausibility would have prevented the erroneous conclusion. If the findings were attributable to the chemicals studied, bilirubin and GGTP would have been increased as well. Enlarged livers and normal means of serum values are certainly not evidence of a toxic effect of chemicals. No information was provided on the serum values for individuals with an enlarged liver; further, there were no data that the livers were truly enlarged. The authors did not disclose the outcome of the medical followup, and no subsequent report was published.

In evaluating VOC toxicity, it should be kept in mind that very little is known about the minimum period of exposure and followup required for an effect to be observed. This is true not only for VOCs but for chemical compounds in general. For instance, experience with known human carcinogens suggests that the period from the first exposure to the development of an observable cancer is usually measured in decades and may be in excess of 30 years (48). Thus, the period of followup in this study was too short and the study size too small to reveal a

positive association, even if the VOCs were carcinogenic to humans. The second factor determining whether an effect of exposure can be observed is the level of exposure. Since no health effects were observed, the exposure levels in this study can be considered a NOEL. Because the exposure levels changed over time, and because of the wide range in reported consumption of water, it is difficult to define a single NOEL characterizing the study. Several NOELs have been calculated instead, as listed below.

1. 13,455 ppb is the *maximum* total VOC concentration in current water samples.
2. 84 ppb is the *median* total VOC concentration in current water samples.
3. 39,367,073 ppb-months is the *maximum TAEVOC* (total accumulated exposure).
4. 4666 ppb-months is the *median* of the positive TAEVOC values.
5. 0.33 mg/day for 159 months is the *minimum* total accumulated dose (DOSEVOC) of the upper 25% of all people with a positive DOSE value, averaged over the duration of exposure.
6. 4.14 mg/day for 60 months is the *median* DOSEVOC of the upper 25% of DOSEVOC values, averaged over the duration of exposure.
7. 477.18 mg/day (at a 70 kg average body weight equal to 6.8 mg/kg/d) for 99 months is the *maximum* DOSEVOC value averaged over the duration of exposure.

The maximum NOEL is approximately the same as found in animal and occupational studies of exposure to single chemicals (Chapter 3).

9.6 Study Findings Unrelated to VOC Exposure

The data compiled in this study allowed for an evaluation of questions other than effects related to VOC exposures. The results of an exploratory analysis of disease risk factors are presented in Appendices E through H. Additional information could probably be extracted if more time were available. Further analyses, most of which would require specifically focused field studies would be of great scientific importance, since epidemiologic information is scarce for most of the diseases analyzed.

The analysis yielded information reflecting other public health problems of unknown extent. These questions are listed below; the list is certainly not exhaustive, and the sequence is entirely arbitrary. Noting these issues does not imply that they reflect purely local health problems. Some or all may represent a general situation, and some may touch upon well-known problems.

9.6.1 Clinical Chemistry

Chapter 8.2 discussed the high prevalence of abnormal clinical chemistry test results, unrelated to exposure to VOCs. The underlying cause may be that the upper limit of normal test values (used at the subcontractor's medical center) had been set too low. There is also the possibility that an unrecognized excess of certain diseases caused the high frequency of elevated values. It would require only a relatively simple, low-cost investigation to find an explanation for the excess of abnormal laboratory values.

9.6.2 Exogenous Chemicals

This study yielded one of the largest extant data sets on DDT, PCB, and PBB. Considering the wealth of information on subject health status and clinical chemistry, and information on residential and occupational exposure to these chemicals, there is a good possibility that further

analysis could lead to significant conclusions. Chlordane has recently raised public concern in Michigan. The finding that only a few people had measurable levels might indicate that they had been exposed while applying pesticides at work or at home. This data set contains some information on such exposure.

9.6.3 Diabetes

Over one quarter of the diabetes cases reported had received no treatment, partially (three out of 28) because they were newly found cases. Why the others were untreated was unknown; some or all might have been on dietary restrictions only. The 1980-1984 diabetes mortality in Battle Creek City was one third higher than the State rate, albeit not statistically significant. This raises the question of whether increased diabetes mortality might be associated with unawareness of the disease, poor compliance with therapy, or difficulties in access to medical care. The high risk of diabetes (odds ratio of 4.1) among people who had a first degree relative with diabetes warrants a further study to analyze the role of the kind of family relation and the type of diabetes, in order to fully understand the public health implications.

9.6.4 Hypertension

About 14% of the persons with hypertension had no medical record or were not on medication, and were unaware they had borderline (6%) or definite (8%) hypertension. Another 14% had medication but were unaware of the disorder, and four out of 92 medicated people clearly had *hypotension* (but did not know they had been taking antihypertensives for up to 5 years). These findings suggest one or more of the following problems: a) insufficient public education in health; b) poor compliance to therapy or lack of medical followup of medicated people, c) inadequate physician-patient communication, and d) financial barriers to access to medical services. In the study population, the degree of awareness of established hypertension is much lower than in the U.S. population, although the percentage of people on pharmacotherapy and the prevalence of hypertension are higher. This raises the question of the possible over-diagnosis of hypertension and, thus, of over-treatment with the associated risk of serious side effects. A closer look is indicated into the usage patterns of antihypertensive medications and the basis for their prescription.

9.6.5 Pregnancy Outcomes (Birth Defects not Included)

The study yielded a prevalence of "any abnormal pregnancy outcome" about 10% above the national rate, a statistically insignificant difference. There was, however, a larger difference between Battle Creek City and the suburban areas which may reflect a more general phenomenon. Existing State-generated data could be analyzed for the association between pregnancy outcomes and urbanization, and compared with international prevalence rates. Such a study might indicate the risk factors to focus upon in future studies. Although some risk factors have already been identified, e.g., smoking and alcohol use, analytical epidemiological studies specifically focused on miscarriage and low birth weight are likely to reveal more risk factors, some of which may have value for prevention strategies.

9.6.6 Gallbladder Disease

In the study area the prevalence of gall stones or chronic inflammation of the gallbladder is about 10 times as high as in a sample of the U.S. population. Obesity is a known risk factor for the disease; another risk factor identified by exploratory analysis is the presence of diabetes.

Both risk factors were more frequent in the study population than expected. The higher prevalence of risk factors was not great enough, however, to explain the observed large excess of cases. A new study, specifically designed for this purpose, would be needed for finding a better explanation of the high rates.

9.6.7 Hypothyroidism

The prevalence of hypothyroidism is many times that in the US-population. As the disease was almost exclusively associated with overweight, and females far outnumbered males, questions may be raised as to the quality of the diagnosis. Thyroid supplements are sometimes prescribed for obesity, which may not be without adverse effects and, given the obesity still present, may not serve the intended purpose. A study of the extent of this drug use or of other potential causes seems indicated. Such causes may include a dietary iodine deficit (easy to remedy) or occupational exposure to yet unidentified chemicals.

9.6.8 Alcohol Abuse

The mortality and morbidity rates for alcohol-related diseases were above average in Calhoun County (Appendix D). The study yielded valuable information on the level and distribution of alcohol consumption, the type of alcohol-containing product consumed, the age at which people started to drink alcohol, and a complete medical file. These data are sufficient for a first-stage analysis of the extent and magnitude of the problem, and may also yield an explanation for the excess of tests with abnormal values for liver enzymes.

9.6.9 Peptic Ulcer

The prevalence of peptic ulcer is twice that of the US population, a finding for which there is no explanation. A related finding is the large number of antacid users who did not have a history of peptic ulcer, although some might have an unrecognized ulcer. The question may be raised whether there is a relation between use of antacids, incidence of peptic ulcer, and occurrence among antacids users of undetected cases of ulcer. Such an association may, for instance, be related to reduced access to medical services for financial reasons. The study population probably has an income and education level below the national average.

9.6.10 Followup of People with Newly Found Disorders

Once the field component of the Battle Creek Health Study was completed, there was no further followup of the study population with regard to their health status. All study participants received a copy of the results of the clinical examination, and were advised to see their physician for an explanation if they had questions.

9.6.11 Disease Incidence

Table 9.2 provides an overview of the absolute frequency of validated diseases that have been analyzed, regardless of when and where the diagnosis was made, and as rates per 1000 population. The rates are neither incidence rates (they are not restricted to a period of time) nor prevalence rates (deceased persons have not been excluded from the numerator and denominator), but represent what the public in general perceives as disease incidence. These

numbers and rates should *not* be used for an analysis of exposure-disease associations. Yet, these numbers may point to differences in disease rates between areas, based on unknown factors.

Table 9.2 may serve also as a quick reference for the disease specific sections above. Overall, the exposed and reference cohorts have approximately equal numbers of diseases which are more frequent in one area than in the other. This would be expected if the basis for cohort formation (VOCs) is unrelated to excess of disease.

9.7 Summary

To investigate the possible health effects of VOC exposure at environmental levels, a complex methodology was developed based on the concept of risk assessment for toxic chemicals. The methodology was developed as if all required data were or could be made available. The real-world scenario was far from this ideal, however, and compromises were necessary. The most important deficiencies in the exposure assessment were: the lack of reliable information on chemical-specific absorption rates for the lungs and the skin, the lack of NOELS based on a standardized set of toxicity parameters, and the lack of monitoring data for private wells. In the health assessment, deficiencies were created by insufficient quality control for the retrieval and abstracting of medical records. It was inevitable, given a retrospective health study, that the quality of historical health information was less than optimal.

Recognizing these difficulties, the study design proved applicable, and the results were satisfactory. This study is the first epidemiologic study designed to assess the effect of exposure to multiple chemicals via multiple routes, and the first designed as a risk assessment. The use of alternative exposure expressions, quality control procedures for exposure and health data, and the application of essential rules for epidemiologic evaluation, compensated for deficiencies in the data and in knowledge of the disease genesis mechanisms. Improvement over conventional study designs was achieved by developing exposure and analytical models, and by validating health data. In the evaluation of the analytical results, Hill's criteria for causality (44) were given priority over p-values. One of these criteria is internal consistency of the results. The use of a variety of analytical models and exposure expressions provided the means for an extensive application of this criterion.

The conclusions of the study are: VOCs have not had observable adverse health effects and, with the exception of diabetes and psoriasis, it is unlikely that a larger study size or a longer period of followup would have produced any other conclusion. Some diseases did occur in excess in the contaminated area, if all cases are counted regardless of eligibility criteria. On the other hand, other diseases and most abnormal clinical chemistry values were more frequent in the reference area. In addition, some diseases were far more frequent in the entire study area than expected from national prevalence data. It is likely that the extent of this study, and the multiple sources of health data resulted in more complete case-finding and, thus, in higher incidence and prevalence rates. For most of the diseases, the incidence rates are the first ever reported. Further studies are required to find the cause(s) underlying the larger than expected prevalence rates, and to evaluate differences between neighborhoods.

References

References

1. Abbott, R. D. Logistic regression in survival analysis. *Am. J. Epidemiol.* 121: 465-471 (1985).
2. Allen-Rowlands, C. F., Castracane, V. D., Hamilton, M. G., and Seifter, J. Effect of polybrominated biphenyls (PBB) on the pituitary-thyroid axis of the rat. *Proc. Soc. Exp. Biol. Med.* 166: 506-514 (1981).
3. Andelman, J., Meyers, S. M., and Wilder, L. C. Volatilization of organic chemicals from indoor uses of water. In: *Chemicals in the Environment* (J.N. Lester, R. Perry, and R.M. Sterritt, Eds.). Selper Ltd, London. 1986. pp. 323-330.
4. Andelman, J. B. Inhalation exposure in the home to volatile organic chemicals of drinking water. *Sci. Tot. Environ.* 47: 443-460 (1985).
5. Bahn, A. K., Mills, J. L., Snyder, P. J., et al. Hypothyroidism in workers exposed to polybrominated biphenyls. *N. Engl. J. Med.* 302: 31-33 (1980).
6. Bailar, J. C. and Ederer, F. Significance factors for the ratio of a Poisson variable to its expectation. *Biometrics* 20: 639-643 (1964).
7. Bellar, T. A., Lichtenberg, J. J., and Kroner, R. C. The occurrence of organohalides in chlorinated drinking water. *J. Am. Water Works Assoc.* 66: 703-706 (1974).
8. Belloc, N. B. Validation of morbidity survey data by comparison with hospital records. *Am. Statist. Assoc. J.* 49: 832-846 (1954).
9. Bierman, J. M., Siegel E., French, F. E., and Simonian K. Analysis of the outcome of all pregnancies in a community, the Kauai pregnancy study. *Am. J. Obst. Gynec.* 91: 37-45 (1965).
10. Boyce, J. D., Jr. and Monson, R. R. Breast cancer in women after repeated fluoroscopic examinations of the chest. *J. Nat. Cancer Inst.* 59: 823-832 (1977).
11. Brennan, M. J. and Swanson, G. M. Cancer Incidence in Metropolitan Detroit 1973-1977. In: *Cancer Incidence in Five Continents*, Vol. 4, (Waterhouse J. et al., eds.). International Agency for Research on Cancer, Lyon. 1982. pp. 344-351.
12. Breslow, N. E. and Day, N. E. The analysis of case-control studies. In: *Statistical Methods in Cancer Research*. Volume 1. IARC Scientific Publications No. 32 (W. Davis, editor). IARC, Lyon. 1980. p. 115.
13. Brown, H. S., Bishop, D. R., and Rowan, C. A. The role of skin absorption as a route of exposure for volatile organic compounds (VOCs) in drinking water. *Am. J. Publ. Health* 75: 479-484 (1984).
14. Buck, W. B., Osweiler, G. D., and Van Gelder, G. A. *Clinical and Diagnostic Veterinary Toxicology*. Kendal Hunt Publishing Co., Dubuque, Iowa, 2nd edition. 1980. pp 191-204.
15. Bull, R. J., Brown, J. M., Meiserhenry, E. A., et al. Enhancement of the hepato-toxicity of chloroform in B6C3F1 mice by corn oil: implication for chloroform carcinogenesis. *Environ. Health Perspect.* 49: 49-58 (1986).

16. California Department of Health Services. *Pregnancy Outcomes in Santa Clara County 1980-1985*. Berkeley, CA, May 1988.
17. Carpenter, C. P. The chronic toxicity of tetrachloroethylene. *J. Ind. Hyg. Toxicol.* 19: 323-326 (1937).
18. Clark, C. S., Meyer, C. R., Gartside, P. S., Majedi, V. A., and Specker, B. An environmental health survey of drinking water contamination by leachate from a pesticide waste dump in Hardeman County, TN. *Arch. Environ. Health* 37: 9-18 (1982).
19. Committee to Coordinate Environmental and Related Programs (CCERP), U.S. Department of Health and Humans Services (DHHS). Risk assessment and risk management of toxic substances. *Report of the CCERP Executive Committee to the Secretary of DHHS*, April 1985.
20. Dallas, C. E., Weir, F. W., Feldman, S., Putcha, L., and Bruckner, J. V. The uptake and disposition of 1,1-dichloroethylene in rats during inhalation exposure. *Toxicol. Appl. Pharmacol.* 68: 140-151 (1983).
21. Doll, R. and Smith, P. G. Comparison between registries: age-standardized rates. In: *Cancer Incidence in Five Continents*, Vol. 4. (Waterhouse J. et al., eds.). IARC, Lyon, 1982: 671-675.
22. Dorsch, M. M., Scragg, R. K. R., McMichael, A. J., Baghurst, P. A., and Dyer, K. F. Congenital malformations and maternal drinking water supply in rural South Australia: A case-control study. *Am. J. Epidemiol.* 119: 473-486 (1984).
23. Elinson, J. and Trussel, R. E. Some factors relating to degree of correspondence for diagnostic information as obtained by household interviews and clinical examinations. *Am. J. Publ. Health* 47: 311-321 (1957).
24. Environmental Protection Agency. National drinking water regulations; volatile organic chemicals. *Fed. Reg.* 49: 25330-24335 (1984).
25. Environmental Protection Agency. National primary drinking water regulations; monitoring for unregulated chemicals. *Fed. Reg.* 51: 25600- 25717 (1987).
26. Environmental Protection Agency. Guidelines for the health risk assessment of chemical mixtures. *Fed. Reg.* 51: 34014-34025 (1986).
27. Environmental Protection Agency. *Draft Criteria Document for 1,1,1-trichloroethane*. Office of Drinking Water, 1984 (NTIS PB84-199520).
28. Environmental Protection Agency. *Health Assessment Document for 1,1,1-trichloroethane, Final Report*. Office of Research and Development, 1984 (NTIS PB84-183565, EPA 600/8-82-003F).
29. Environmental Protection Agency. *Draft Criteria Document for Tetrachloroethylene*. Office of Drinking Water, 1984 (NTIS PB 84-199512).
30. Environmental Protection Agency. *Health Assessment Document for Tetrachloroethylene (Perchloroethylene), Final Report*. Office of Research and Development, 1985 (NTIS PB85-249704, EPA 600/8-82-005F).

31. Environmental Protection Agency. *Updated Carcinogenicity Assessment for Tetrachloroethylene, Draft Report: Addendum to the Health Assessment Document for Tetrachloroethylene (Perchloroethylene), Final Report.* Off. Research & Development, 1986 (NTIS PB86-174489, EPA 600/8-82-005F).
32. Environmental Protection Agency. *Draft Criteria Document for Trichloroethylene.* Office of Drinking Water, 1984 (NTIS PB 84-209840).
33. Environmental Protection Agency. *Health Assessment Document for Trichloroethylene, Final Report.* Office of Research and Development, 1985 (NTIS PB85-249696, EPA 600/8-82-006F).
34. Environmental Protection Agency. *Draft Criteria Document for 1,2-dichloroethane.* Office of Drinking Water, 1984 (NTIS PB 84-209816).
35. Environmental Protection Agency. *Health Assessment Document for 1,2-dichloroethane, Final Report.* Office of Research and Development, 1985 (NTIS PB86-122702, EPA 600/8-84-006F).
36. Environmental Protection Agency. *Draft Criteria Document for Dichloroethylenes.* Office of Drinking Water, 1984 (NTIS PB 84-199546).
37. Environmental Protection Agency. *Health Assessment Document for Vinylidene Chloride, Final Report.* Office of Research and Development, 1984 (NTIS PB86-100641, EPA 600/8-82-031F).
38. Environmental Protection Agency. *Ambient Water Quality Criteria for Chloroform.* Office of Water Regulations and Standards, 1980 (EPA 440/5-8- 033).
39. Finklea, J., Priester, L. E., Creason, J. P., et al. Polychlorinated biphenyl residues in human plasma expose a major urban pollution problem. *Am. J. Publ. Health* 62: 645-651 (1972).
40. Gordis, L. Assuring the quality of questionnaire data in epidemiologic research. *Am. J. Epidemiol.* 109: 21-24 (1979).
41. Harrell, F. The PHGLM procedure. In: *SUGI Supplemental Library User's Guide*, 5th ed., SAS Institute, Cary, N.C., 1986. pp. 437-466.
42. Herman, W. H., Teutsch, S. M., and Geiss, L. S. (prepared for The Carter Center). Closing the gap: the problem of diabetes mellitus in the United States. *Diabetes Care* 8: 391-406 (1985).
43. Hildreth, N. G., Shore, R. E., Hempelmann, L. H., and Rosenstein, M. Risk of extrathyroid tumors following radiation treatment in infancy for thymic enlargement. *Radiation Research* 102: 378-391 (1985).
44. Hill, A. B. The environment and disease: association or causation? *Proc. Royal Soc. Med.* 58: 295-300 (1965).
45. Hoffman, D. A. Breast cancer risk in scoliotic women exposed to multiple diagnostic X-rays. *Proceedings 22nd Annual Meeting Scoliosis Research Society*, Vancouver, Canada, September 16, 1987.
46. Hofmann, H. Th., Birnstiel, H., and Jobst, P. Zur Inhalationtoxizität von 1,1- und 1,2-Dichloroathan. *Arch. Toxikol.* 27: 248-265 (1971).

47. International Agency for Research on Cancer (IARC). *Some Anti-thyroid and Related Substances, Nitrofurans, and Industrial Chemicals*. Monograph series, Vol. 7. IARC, Geneva, 1974.
48. International Agency for Research on Cancer (IARC). *Some Halogenated Hydrocarbons and Pesticide Exposures*. Monograph series, Vol. 41, IARC, Lyon, 1986, p.22.
49. Kay, J. Polybrominated biphenyls (PBB) environmental contamination in Michigan. *Environ. Res.* 13: 74-93 (1977).
50. Kelsey, J. L., Thompson, W. D., and Evans, A. S. Retrospective cohort studies. In: *Methods in Observational Epidemiology*. Oxford University Press, New York, 1986. pp. 128-147.
51. Kimbrough, R., Buckley, J., Fishbein, L., et al. Animal toxicology. *Environ. Health Perspect.* 24: 173-184 (1978).
52. Kimmerle, G. and Eben, A. Metabolism, excretion, and toxicology of trichloroethylene after inhalation, I. Experimental exposure on rats. *Arch. Toxicol.* 30: 115-126 (1973).
53. Kramer, C. G., Ott, M. C., Fulkerson, J. E., and Hicks, N. Health of workers exposed to 1,1,1-trichloroethane: a matched-pair study. *Arch. Environ. Health* 33: 331-342 (1978).
54. Kreiss, K., Roberts, C., and Humphrey, H. E. B. Serial PBB levels, PCB levels, and clinical chemistries in Michigan's PBB cohort. *Arch. Environ. Health* 37: 141-147 (1982).
55. Lagakos, S. W., Wessen, B. J., and Zelen, M. An analysis of contaminated well water and health effects in Woburn, Massachusetts. *J. Am. Stat. Assoc.* 81: 583-596 (1986).
56. Lawrence, C. E. and Taylor, P. R. Empirical estimation of exposure in retrospective epidemiologic studies. In: *Environmental Epidemiology* (F. C. Kopfler and G. F. Craun, Eds.). Lewis Publishers, Chelsea, Michigan, pp. 239-246, 1986.
57. Liang, K., and Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika* 73: 13-22 (1986).
58. MacMahon, B. Comment (Ref. 31). *J. Am. Stat. Assoc.* 81: 597-599 (1986).
59. Maltoni, C., Lefemine, G., Chieco, P., Cotti, G., and Patella, V. Experimental research on vinylidene chloride carcinogenesis. In: *Archives of Research on Industrial Carcinogenesis*, volume 3, (Maltoni C. and Mehlman M.A., Eds.), Princeton Scientific Publications, Princeton NJ, 1985.
60. Maltoni, C., Valgimigli, L., and Scarnato, C. Long term carcinogenic bioassays on ethylene-dichloride administered by inhalation to rats and mice. In: *Ethylene Dichloride - A Potential Health Risk?* Banbury Report 5 (B. Ames, P. Infante, and R. Reitz, Eds.), Lloyd Harbor, NY, 1980. pp. 3-32.
61. McKone, T. E. *Quantifying Personal Indoor Air Exposure to Radon and Volatile Organic Compounds Derived from Tap Water*. Annual Meeting of the Society for Risk Analysis, Houston, Texas, November 1987.
62. Moslen, M. T., Reynolds, E. S., and Szabo, S. Enhancement of the metabolism and hepatotoxicity of trichloroethylene and perchloroethylene. *Biochem. Pharmacol.* 26: 369-375 (1977).

63. Murphy, P. J. *The Great Oaks Water Distribution Study*. Report to the California Department of Health Services, April 1988.
64. National Cancer Institute. *Bioassay of 1,1,1-trichloroethane for Possible Carcinogenicity*. DHEW Publ. no. (NIH) 77-803, NCI-CG-TR- 3, Washington, DC, 1977.
65. National Cancer Institute. *Bioassay of 1,2-dichloroethane for Possible Carcinogenicity*. DHEW Publ. no. (NIH) 77-1361, NCI-CG-TR-55, Washington DC, 1978.
66. National Cancer Institute. *Carcinogenesis Bioassay of Trichloroethylene*. DHEW Publ. No. (NIH) 76-802, NCI-CG-TR-2, Washington DC, 1976.
67. National Cancer Institute. *Bioassay of Tetrachloroethylene for Possible Carcinogenicity*. DHEW Publ. No. (NIH) 77-813, NCI-CG-TR-13, Washington, DC, 1977.
68. National Center for Health Statistics. *Blood Pressure Levels in Persons 18-74 Years of Age in 1976-1980, and Trends in Blood Pressure from 1960-1980 in the United States*. Data from the National Health Survey, Series 11, No. 234. DHHS Publ. No. (PHS) 86-1684, Hyattsville, MD, 1986.
69. National Center for Health Statistics. *Prevalence of Selected Chronic Conditions United States 1979-1981*. Data from the National Health Survey Series 10, No. 155. DHHS Publ. (PHS) 86-1583, Hyattsville, MD, 1986.
70. National Center for Health Statistics. *Prevalence of Diagnosed Diabetes, Undiagnosed Diabetes, and Impaired Glucose Intolerance in Adults 20-74 Years of Age United States, 1979-81*. Data from the National Health Survey, Series 11, No. 237. DHHS Publ. No. (PHS) 86-1583, Hyattsville, MD, 1987.
71. National Research Council. *Chloroform, Carbon Tetrachloride, and Other Halomethanes: An Environmental Assessment*. Natl. Academy of Sciences, Washington, DC, 1978.
72. National Toxicology Program. *Carcinogenesis Bioassay of 1,1,1-trichloroethane in F344 Rats and B6C3F1 Mice (Gavage Studies)*. Draft report 1983.
73. National Toxicology Program. *Carcinogenesis Bioassay of Vinylidene Chloride in F344 Rats and B6C3F1 Mice (Gavage Studies)*. DHHS Publ. No. (NIH) 77-1784, NTP-TR-228, Washington, DC, 1982.
74. National Toxicology Program. *Toxicology and Carcinogenesis Studies of Tetrachloroethylene (Perchloroethylene) on F344/N Rats and B6C3F1 Mice (Inhalation Studies)*. DHHS Publ. No. (NIH) 86-2567, NTP-TR-311, Research Triangle Park, NC, 1986.
75. National Toxicology Program. *Toxicology and Carcinogenesis Studies of Chlorodibromomethane in F344/N Rats and B6C3F1 Mice (Gavage Studies)*. DHHS Publ. No. (NIH) 85-2538, NTP-TR-282, Research Triangle Park NC, 1985.
76. National Toxicology Program. *Toxicology and Carcinogenesis Studies of Trichloroethylene in Four Strains of Rats (Gavage Studies)*. DHHS Publ. No. (NIH) 88-2529, NTP-TR-273, Research Triangle Park NC, 1988.
77. National Toxicology Program. *Quarterly Program Status Report on Superfund Activities (January 1 to March 31, 1985)*. Research Triangle Park, NC, 1985.

78. Office of Science and Technology Policy. Chemical carcinogens: a review of the science and its associated principles. *Federal Register* 1985, pp. 10371-10442.
79. Office of Technology Assessment, U.S. Congress. *Environmental Contaminants in Food*. Washington, DC, December 1979.
80. Palmer, A. K., Street, A. E., Roc, F. J. C., et al. Safety evaluation of toothpaste containing chloroform, part II long-term studies in rats. *J. Environ. Pathol. Toxicol.* 2: 821-833 (1979).
81. Quast, J. F., Rampy, L. W., Balmer, B. D. J., and Gehring, P. J. *Toxicological and Carcinogenic Evaluation of 1,1,1-trichloroethane Formulation by Chronic Inhalation in Rats*. Unpublished report, 1978.
82. Quast, J. F., Humiston, C. G., Wade, C. E., et al. A chronic toxicity and oncogenicity study in rats and subchronic toxicity study in dogs on ingested vinylidene chloride. *Fundam. Appl. Toxicol.*, 3: 55-62 (1983).
83. Rothman, K. J. Types of epidemiologic studies. In: *Modern Epidemiology*. Little, Brown, & Co., Boston, 1986; pp. 51-76.
84. Seki Y., Urashima Y., Aikawa H., et al. Trichloro compounds in the urine of humans exposed to methylchloroform at subthreshold levels. *Int. Arch. Arbeitsmed.* 34: 39-49 (1975).
85. Selikoff, I. J. and Anderson, H. A. *Health Effects of Polybrominated Biphenyl Exposure in the General Population of Michigan*. Quarterly report to the Michigan Department of Public Health, January - March 1979, Mount Sinai Environmental Science Laboratory, New York, NY.
86. Smyth, H. F. and Carpenter, C. P. The place of the range finding test in the industrial laboratory. *J. Ind. Hyg. Tax.* 26: 269-273 (1944).
87. Smyth, H. F. and Carpenter, C. P. Further experience with the range finding test in the industrial laboratory. *J. Ind. Hyg. Tax.*, 30: 63-68 (1948).
88. Smyth, H. F., Carpenter, C. P., Weil, C. S., et al. Range-finding toxicity data: list VI. *Am. Ind. Hyg. Assoc. J.* 23: 95-107 (1962).
89. Smyth, H. F., Carpenter, C. P., Weil, C. S., et al. Range-finding toxicity data: list VII. *Am. Ind. Hyg. Assoc. J.* 30: 470-476 (1969).
90. Steinberg, K. K., Freni - Titulaer, L. W. J., Rogers, T. N., et al. Effects of polychlorinated biphenyls and lipemia on serum analytes. *J. Toxicol. Environ. Health* 19: 369-381 (1986).
91. Takamatsu, M. Health hazards in workers exposed to trichloroethylene vapor, II. *Kumamoto Med. J.* 15: 43-54 (1962).
92. TNO (Netherlands Organization for Applied Scientific Research). *Toxicology of 1,1-dichloroethane*. Rijswijk, Netherlands, 1981 (NTIS PB 82-182809).
93. Trussell, R. E., Elinson, J., and Levin, M. L. Comparisons of various methods of estimating the prevalence of chronic disease in a community -- the Hunterdon County Study. *Am. J. Publ. Health* 46: 173-182 (1956).

94. Tuttle, T. C., Wood, G. D., Grether, C. B., Johnson, B. L., and Xintaras, C. *A Behavioral and Neurological Evaluation of Dry Cleaners Exposed to Perchloroethylene*. DHEW (NIOSH) Publication 77-214, 1977 (NTIS PB 274794).
95. Wright, B. M. and Dore, C. F. A random-zero sphygmomanometer. *Lancet* 1: 337-338 (1970).
96. Zeger, S. L., Liang, K., and Albert, P. S. *Model for Longitudinal Data: A Generalized Estimating Equation Approach*. Dept. of Biostatistics, School of Public Health, Johns Hopkins University, 1987.
97. Zeise, L., Wilson, R., Crouch, E. Use of acute toxicity to estimate carcinogenic risk. *Risk Analysis* 4: 187-270 (1984).

Appendix A

The Questionnaire Used in Phase II of the Study

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

and

**Arthur W. Bloomer, MS
Center for Environmental Health Sciences
Michigan Department of Public Health
Lansing, Michigan**

ID Number _____

Comments _____

ID Number from Phase I Study _____

Interviewer Name _____ Interviewer
Code Number _____

Month Day Year
Date of Interview _____

Hour Minute
Time Interview Began _____ (24 hr. basis)

Designated Medical Research Project
"Epidemiologic Study of the Potential Health
Effects of Chronic Exposure to Drinking
Water Contaminated With Chlorinated
Short-Chain Aliphatic Hydrocarbons"

Phase II

Hello. I'm (YOUR NAME) from the Kercher Center for Social Research at Western Michigan University. A member of our staff contacted you to arrange this interview appointment. As you know, we are interviewing households in the (Verona; Springfield; Dowagiac) area that were exposed to contaminated drinking water, and we are also interviewing some households from areas without water contamination for purposes of comparison.

As a participant in the second phase of research, we would like to ask some questions about yourself, your use of water, your exposure to chemicals at work and at home, and your medical history. Later a member of our staff will phone you to set up an appointment for you to provide a blood and urine sample and to have your blood pressure, height, and weight measured. All the information you provide will be used to help determine the health effects of exposure to contaminated drinking water, and all of the information will be kept confidential.

If you wish to participate in the second phase of the study, I will need to have your signature on the Statement of Informed Consent that says you agree to take part in the study and on the Patient Information Release Authorization to give us permission to contact physicians and/or hospitals to get additional information on your health history. If you have further questions about the study, I will try to answer them or you may call the person in charge of the data collection -- Dr. James C. Petersen at Western Michigan University (383-1458).

8. What was your (subject's) address before you (subject) moved here?

9a. Was the water supply at this former address contaminated by chemicals?
____ Yes
____ No
____ Don't know

9b. Was your water supply tested for contamination?
____ Yes
____ No
____ Don't Know

9c. (IF YES) Who tested it? _____

9d. (IF CONTAMINATED--9a), What chemicals were involved?

9e. When did this contamination began? Year ___ __

10. What is the date of (your/subject's) birth? Month Day Year

(IF SUBJECT IS DECEASED) Month Day Year
When did (subject) die? _____

Where did (his/her) death occur? State: _____
County: _____

11. What is (your/subject's) Social Security Number? _____

12. (DO NOT ASK RESPONDENT) Sex: ___ Male
___ Female

13. (DO NOT ASK RESPONDENT) Race: ___ White (not Hispanic)
___ White, Hispanic
___ Black
___ Other

14. How many years of education (have you/has subject) completed?
____ Elementary school
____ Some high school
____ High school graduate
____ Some college
____ College graduate
____ Graduate degree
____ Don't know

WATER USE

15. (FOR EXPOSED SUBJECTS ONLY - Verona, Springfield, Dowagiac; if subject was TRACKED, questions 15a + b refer to former residence in study area)

a. When was your residence connected to the municipal water system?

Month Year

____ Has not been connected

b. When did you stop drinking private well water?

Month Year

16a. What is the household's **current** source of drinking water?

____ public system

____ private well or spring

____ bottled water

____ other Specify: _____

____ don't know

16b. When did your household begin to use this source of drinking water?

Month Year

17. Before you (subject) began using the current water source, what previous water sources did you (subject) use? If possible, we could like to have this information back to 1970.

From		To		Public water	Water Source	
Month	Year	Month	Year		Private well	Other
____	____	____	____	____	____	____
____	____	____	____	____	____	____
____	____	____	____	____	____	____
____	____	____	____	____	____	____
____	____	____	____	____	____	____
____	____	____	____	____	____	____

Now I would like to ask you several questions about how much water you consume on an average day. In each case, I would like to know how much you consume at your residence and also how much you consume on the job _____ away from the home.

18. About how many 8 ounce glasses/cups of plain tapwater (do you/does subject) drink on an average day?

At Residence

____ 8 ounce glasses/cups

____ don't know

On The Job

____ 8 ounce glasses/cups

____ don't know

19. How much plain bottled water (do you/does subject) drink on an average day?

At Residence

____ 8 ounce glasses/cups

____ don't know

On The Job

____ 8 ounce glasses/cups

____ don't know

20. How much water used to prepare a powdered or frozen cold drink (do you/does subject) drink on an average day?
 _____ 8 ounce glasses/cups _____ 8 ounce glasses/cups
 _____ don't know _____ don't know
21. How much water used after heating or boiling--such as for coffee or tea--(do you/does subject) drink on an average day?
 _____ 8 ounce glasses/cups _____ 8 ounce glasses/cups
 _____ don't know _____ don't know

Another means of exposure to water in the home is through showers or baths.

- 22a. During the summer, about how many times a week (do you/does subject) take a shower?
 _____ don't know
- 22b. During the winter, about how many times a week (do you/does subject) take shower?
 _____ don't know

IF TAKES SHOWERS

- 22c. When you take a shower, about how many minutes (do you/does subject) spend in the shower?
 _____ minutes
 _____ don't know
- 22d. During the summer, about how many times a week (do you/does subject) take a bath?
 _____ don't know
- 22e. During the winter, about how many times a week (do you/does subject) take a bath?
 _____ don't know

IF TAKES BATHS

- 22f. When you take a bath, about how many minutes (do you/does subject) spend in bathtub?
 _____ minutes
 _____ don't know
23. About how many times a month (do you/does subject) swim in a home pool?
 _____ not applicable
 _____ don't know

(IF THERE IS A HOME POOL)

23b. How many months a year (do you/does subject) use the pool?

- not applicable
- don't know

CHEMICAL EXPOSURE

24. In the home, (have you/has subject), because of a hobby (such as furniture refinishing, woodworking, photo developing, painting, or auto repair), been exposed to any of the following substances:

- a. Solvents (such as turpentine, degreasers, mineral spirits, brush cleaners, etc.)
 - Yes Specify: _____
 - No
 - Don't know
 - No answer
- b. Paints
 - Yes Specify: _____
 - No
 - Don't know
 - No answer
- c. Glues or resins
 - Yes Specify: _____
 - No
 - Don't know
 - No answer

25. Does this household **currently** subscribe to any type of pest control service for inside the home?

- Yes (IF YES) Can you tell me the chemical used or the name of the company? _____
- No
- Don't know
- No answer

26. Does this household **currently** subscribe to any chemical services for the yard or outdoor part of the home (such as chemical treatment of the lawn)?

- Yes (IF YES) Can you tell me the chemical used or the name of the company? _____
- No
- Don't know
- No answer

27. Does this household **regularly** use any insecticide or anti-termite chemical inside or outside the house?

- Yes (IF YES) Can you tell me the chemical used or the name of the company? _____
- No
- Don't know
- No answer

28. Does this household regularly use any weed killers or other such chemicals under or outside the house?

- Yes (IF YES) Can you tell me the chemical used or the name of the company? _____
- No
- Don't know
- No answer

28b. On the average, how many times a week do you eat fresh fish purchased from a grocery store or fish market? _____

28c. On the average, how many times a week do you eat fresh fish that your family or friends have caught in Michigan waters? _____

Now I would like to ask you a few questions about possible exposure to chemicals at work. First I would like to know what kind of work (you/subject) currently do and also any other jobs (you/subject) may have held since 1970.

29. Current Job

From (Year)	To (Year)	Occupation	Describe Duties
---	---	_____	_____

29b. Previous Jobs (Start With Most Recent)

From (Year)	To (Year)	Occupation	Describe Duties
---	---	_____	_____
---	---	_____	_____
---	---	_____	_____
---	---	_____	_____

30. (Have you/has subject) been exposed at work to chemicals such as solvents, degreasers, dyes, pesticides, chemicals in the rubber and plastic industry, glues and resins, etc.?

- yes
- No
- Don't know
- No answer

(IF YES)

30b. For each chemical exposure, I would like to know the chemicals involved, the years of exposure, the average number of hours of exposure per day, and whether or not (you/subject) used protective measures such as protective clothing, gloves, or masks.

Chemical	From To		Hrs of exposure per day	Protective measures used			
	year	year		Yes	No	DK	NA
_____	---	---	---	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_____	---	---	---	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_____	---	---	---	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_____	---	---	---	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
_____	---	---	---	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

31. (Have you/has subject) ever been employed in a dry cleaning facility?

- Yes
- No
- Don't know
- No answer

(IF YES).

31b. When (were you/was subject) employed there and what duties did you have?

From (Year) To (Year) Duties

LIFESTYLE HABITS

Now I would like to ask you a few brief questions about your lifestyle -- about your use of alcohol or tobacco.

32. Did (you/subject) ever drink alcoholic beverages?

- Yes (IF YES, CONTINUE)
- No (IF NO, GO TO QUESTION 36)
- Don't know
- No answer

33. At what age did (you/subject) begin to drink?

- (Years of Age)
- Don't know
- No answer

34. Do (you/subject) now drink alcoholic beverages?

- Yes
- No
- Don't know
- No answer

(IF NO)

34b. How many years ago did you quit drinking alcoholic beverages?

- Years
- Don't know
- No answer

35. On the average, how much (do/did) (you/subject) drink of the following?

- 35b. Beer 12 oz. glasses/12 oz. cans per day/week/month
- 35c. Wine glasses/bottles per day/week/month
(IF BOTTLES ask size)
- 35d. Whiskey shots/bottles per day/week/month
(IF BOTTLES ask size)
- 35e. Other (specify) _____ per day/week/month
(IF BOTTLES ask size)
- 35f. _____ per day/week/month
(IF BOTTLES ask size)

36. Did (you/subject) ever smoke or use tobacco products?

- Yes (IF YES, CONTINUE)
- No (IF NO, GO TO QUESTION 40)
- Don't know
- No answer

37. At what age did (you/subject) begin smoking?

- ___ (years of age)
- Don't know
- No answer

38. Do (you/subject) now smoke or use tobacco products?

- Yes
- No
- Don't know
- No answer

38b. How many years ago did you quit?

- ___ (years)
- Don't know
- No answer

39. On the average, how much (do/did) (you/subject) use of the following?

- 39b. Cigarettes ___ cigarettes/packs per day
- 39c. Cigars ___ cigars per/day
- 39d. Cigarillos ___ cigarillos/packs per day
- 39e. Chewing tobacco ___ ounces per day
- 39f. Pipes ___ bowls per day

MEDICAL HISTORY

40a. Now I would like to ask you a number of questions about your medical history. First, For what diseases or illnesses (are you/is subject) currently receiving medical care or treatment? Who is treating you for the condition?

Condition	Name & Address of Physician
_____	_____
_____	_____
_____	_____
_____	_____

40b. (Do you/Does subject) now take any medications prescribed by a physician? (Include insulin, drugs used to control hypertension, and painkillers and sleeping pills if regularly used.) HAND SUBJECT CARD. These are some common types of medications.

- Yes
- No
- Don't know
- No answer

ID Number _____

(IF YES)

40c. For each medication now taken, I would like to know its name, when (you/subject) began taking it, and the name and address of the physician who prescribed it.

(ASK RESPONDENT TO READ INFORMATION FROM LABELS OF BOTTLES)

Medication	Taken since		Name and address of physician
	Month	Year	
_____	__	__	_____
_____	__	__	_____
_____	__	__	_____
_____	__	__	_____

41. (Do you/Does subject) now take any "over the counter" medications - that is, remedies that can be purchased without a prescription? HAND RESPONDENT CARD. These are some common over the counter medications.

- Yes
- No
- Don't know
- No answer

(IF YES)

41b. For each over the counter medication you now take, I would like to know its name and when (you/subject) began taking it.

OTC Medication	Taken since	
	Month	Year
_____	__	__
_____	__	__
_____	__	__
_____	__	__

ID Number _____

42. Since 1970, (have you/has subject) been seen by a doctor for any of the following conditions?

(IF YES ON QUESTIONS 42-46), ASK: When (were you/was subject) first seen for this condition?
Who was the attending physician? (GET ADDRESS OF PHYSICIAN).
Were you hospitalized? (IF YES, GET NAME AND ADDRESS OF HOSPITAL.)

First, for various skin conditions:

	Date first seen monthyear	Name & address Attending Physician	Hospitalized	Name & address of Hospital	Seen before 1970?	Year
42a. Psoriasis						
___ Yes ___ No	___ ___	_____	___ Yes ___ No	_____	___ Yes ___ No	___
		_____		_____		
42b. Dermatitis						
___ Yes ___ No	___ ___	_____	___ Yes ___ No	_____	___ Yes ___ No	___
		_____		_____		
42c. Eczema						
___ Yes ___ No	___ ___	_____	___ Yes ___ No	_____	___ Yes ___ No	___
		_____		_____		
42d. Acne						
___ Yes ___ No	___ ___	_____	___ Yes ___ No	_____	___ Yes ___ No	___
		_____		_____		
42e. Hives						
___ Yes ___ No	___ ___	_____	___ Yes ___ No	_____	___ Yes ___ No	___
		_____		_____		
42f. Other (specify)						
_____	___ ___	_____	___ Yes ___ No	_____	___ Yes ___ No	___
		_____		_____		

ID Number _____

43. Since 1970, (have you/has subject) been seen by a doctor for any of the following liver problems?

	Date first seen monthyear	Name & address Attending Physician	Hospitalized	Name & address of Hospital	Seen before 1970?	Year
43a. Hepatitis			Yes ___ No ___		Yes ___ No ___	
43b. Cirrhosis			Yes ___ No ___		Yes ___ No ___	
43c. Gall bladder disease			Yes ___ No ___		Yes ___ No ___	
43d. Yellow jaundice			Yes ___ No ___		Yes ___ No ___	
43d. Other (specify)			Yes ___ No ___		Yes ___ No ___	

Blood and vascular disorders: How about blood and vascular disorders?

	Date first seen monthyear	Name & address Attending Physician	Hospitalized	Name & address of Hospital	Seen before 1970?	Year
43f. Anemia			Yes ___ No ___		Yes ___ No ___	
43g. Leukemia			Yes ___ No ___		Yes ___ No ___	

ID Number _____

43h. Hypertension
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____ Year

43i. Other (specify)
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44. Urinary/Renal: Since 1970, (have you/has subject) been seen by a doctor for any of the following urinary or renal problems?

	Date first seen month year	Name & address Attending Physician	Hospitalized	Name & address of Hospital	Seen before 1970?	Year
--	-------------------------------	---------------------------------------	--------------	-------------------------------	-------------------	------

44a. Nephritis
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44b. Renal (kidney) failure
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44c. Urinary tract infection
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44d. Kidney stones
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44e. Protein in the urine
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44f. Blood in the urine
___ Yes ___ No _____ Yes ___ No _____ Yes ___ No _____

44g. Other kidney disease (specify)
_____ Yes ___ No _____ Yes ___ No _____

ID Number _____

Nervous system: How about problems of the nervous system?

	Date first seen monthyear	Name & address Attending Physician	Hospitalized		Name & address of Hospital	Seen before 1970?		Year
			Yes	No		Yes	No	
44. Seizures, fits, epilepsy, convulsions	___ Yes ___ No ___	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	_____
44i. Stroke	___ Yes ___ No ___	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	_____
44j. Weakness or paralysis in arms or legs	___ Yes ___ No ___	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	_____
44k. Emotional problems	___ Yes ___ No ___	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	_____
44l. Other (specify)	_____	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	_____

ID Number _____

45. Metabolic conditions: Since 1970, (have you/has subject) been see by a doctor for any of the following conditions ?

	Date first seen monthyear	Name & address Attending Physician	Hospitalized	Name & address of Hospital	Seen before 1970?	Year
45a. Diabetes						
	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____		_____		

45b. (IF YES) (Do you/Does subject) take insulin to treat the diabetes ?
 Yes No Don't know NA

45c. Drinking problems	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____		_____		

Allergic Conditions

45d. Asthma	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____		_____		

45e. Hay Fever	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____		_____		

45f. Other allergic conditions (specify)		_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No	_____
		_____		_____		

ID Number _____

Cancer

	Date first seen		Name & address Attending Physician	Hospitalized		Name & address of Hospital	Seen before 1970?		Year
	month	year		Yes	No		Yes	No	
46. Any type of cancer	_____	_____	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	
46a.	_____	_____	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	
46b.	_____	_____	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	
46c.	_____	_____	_____	___ Yes ___ No	_____	_____	___ Yes ___ No	_____	

47. (Have you/Has subject often or frequently noticed any of the following symptoms?)

47a. Headaches	___ Yes ___	No ___	Don't know ___	NA ___
47b. Fatigue	___ Yes ___	No ___	Don't know ___	NA ___
47c. Nausea	___ Yes ___	No ___	Don't know ___	NA ___
47d. Sleepiness	___ Yes ___	No ___	Don't know ___	NA ___
47e. Nervousness	___ Yes ___	No ___	Don't know ___	NA ___
47f. Malaise	___ Yes ___	No ___	Don't know ___	NA ___
47g. Irritability	___ Yes ___	No ___	Don't know ___	NA ___

ID Number _____

48. Since 1970, (have you/has subject) had any other diseases or physical or emotional condition not mentioned above ?

IF YES, ASK: What was the disease? When (were you/was subject) first seen for the condition?

(GET NAME AND ADDRESSES OF PHYSICIAN AND HOSPITAL) (ASK: WERE YOU SEEN FOR THIS BEFORE 1970? ASK YEAR

Disease (specify)	Date first seen monthyear	Name & address Attending Physician	Hospitalized		Name & address of Hospital	Seen before 1970?		Year
			Yes	No		Yes	No	
48a.	_____	_____	Yes	No	_____	Yes	No	_____
48b.	_____	_____	Yes	No	_____	Yes	No	_____
48c.	_____	_____	Yes	No	_____	Yes	No	_____

Family history

49. Have any of (your/subject's) immediate family (parents, brothers or sisters, children) ever had diabetes?

___ Yes ___ No ___ Don't know ___ NA

(IF YES) Which family members had diabetes? Was insulin used to treat the diabetes?

Family member (relation) Insulin used ?

49b. _____ Yes ___ No ___ Don't know ___ NA

49c. _____ Yes ___ No ___ Don't know ___ NA

49d. _____ Yes ___ No ___ Don't know ___ NA

ID Number _____

50. Have any of (your/subject's) immediate family (parents, brothers or sisters, children) ever had cancer?
___ Yes ___ No ___ Don't know ___ NA

(IF YES) Which family members had diabetes? Was insulin used to treat the diabetes?
Family member (relation) Type of cancer (site)

- 50b. _____
50c. _____
50d. _____
50e. _____

51. Has anyone in (your/subject's) immediate family had liver problems?
___ Yes ___ No ___ Don't know ___ NA
52. Has anyone in (your/subject's) immediate family had yellow jaundice?
___ Yes ___ No ___ Don't know ___ NA

FOR MALE SUBJECTS - TERMINATE INTERVIEW

Thank you for your cooperation in completing this health survey. The information you have provided will be very helpful to the researchers in this project.

53. (Have you/Has subject) ever been pregnant ?
___ Yes ___ No ___ Don't know ___ NA

IF YES, CONTINUE
IF NO, GO TO 56

ID Number _____

54. I would like to collect some information on each of your pregnancies, starting with the first one. I would like to know if the pregnancy ended with a full term live birth, a stillbirth, or a miscarriage; when this occurred; and at what hospital. FOR FULL TERM, PREMATURE, STILLBORN ASK: What was the birth weight? Were any birth defects present?

Preg-nancy	Outcome	Date	B. Weight	Defect	Hospital							
						Full term live	Prema-ture live	Still-born	Miscar-riage or sp.abort.	Mo Yr	Pound Oz.	Present (specify)
54a.	1	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54b.	2	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54c.	3	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54d.	4	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54e.	5	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54f.	6	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54g.	7	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54h.	8	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54i.	9	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____
54j.	10	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____	_____

55. Since 1970, (have you/has subject) had a child that died before his/her first birthday ?
 ___ Yes ___ No ___ Don't know ___ NA

(IF YES) May I have the child's name, the dates of birth and death, and the name of the hospital?

Name	Date of birth			Date of death			Hospital	
	Mo	Day	Year	Mo	Day	Year	Name	Address
55b	_____	_____	_____	_____	_____	_____	_____	_____
	_____	_____	_____	_____	_____	_____	_____	_____
55c	_____	_____	_____	_____	_____	_____	_____	_____
	_____	_____	_____	_____	_____	_____	_____	_____

56. (Do you/Does subject) currently use oral contraceptives ?
 ___ Yes ___ No ___ Don't know ___ NA

Thank you for your cooperation in completing this health survey. The information you have provided will be very helpful to the researchers in this project.

Time interview concluded: ___ hour ___ minute (24 hr. basis)

INTERVIEWER: Record any evaluations of the subject or comments about the interview.

Appendix B

**Report on the Field Component of the
Study of Health Effects From Contaminated
Drinking Water in Battle Creek and Dowagiac**

**James C. Petersen, PhD
Principal Investigator
Kercher Center for Social Research
Western Michigan University
Kalamazoo, Michigan**

B.1 Introduction

The Contract Agreement 3036 between the State of Michigan and the Leonard Kercher Center for Social Research at Western Michigan University provided for Phase II of a study of the health effects of exposure to short-chain chlorinated hydrocarbons. During the period of June 1, 1985 through August 31, 1986, the Kercher Center, in cooperation with the Michigan Department of Public Health (MDPH) and the Centers for Disease Control (CDC), conducted a study of these health effects within research sites located in Calhoun County and Cass County. Prior to the current contract, the Kercher Center for Social Research had conducted interviews (primarily by telephone) in Phase I of the research (April-September, 1984). In addition to interviewing households identified as "exposed" by the Michigan Department of Public Health, approximately five potential control households were identified for each exposed case. In this initial phase, one member per household was interviewed. As part of that interview, the respondents from both exposed families and potential comparison families were asked if they would be willing to participate in the more detailed Phase II study.

During Phase II of the research, attempts were made to gain the participation of each eligible member of the exposed households and each eligible member of comparison households selected by the CDC. For a person to be eligible to participate in Phase II, he or she must have lived at the address included in the study for at least twelve consecutive months between January 1, 1975 and January 1, 1984.

The Phase II personal interviews with participants in the research began on July 11, 1985. Potential participants were first contacted by letter to explain the project and provide them in advance with a copy of the informed consent form and the medical records release that were used in the project. They then received a telephone call from a member of the research project staff to answer any questions about the research project, ask if they were willing to participate, and schedule an interview in the participant's home. Appointments for the collection of blood and urine samples and physical measurements began on August 19, 1985. The preparation of computer software for the abstracting of physician and hospital records began in July, 1985. The initiation of actual abstracting of records occurred in August, 1985.

In the initial phone contact with the family to be interviewed, the responding adult was asked the year in which each family member first lived at the address of interest. In the case of tracked interviews, the year in which each family member moved out of the address of interest was obtained. When the eligibility was clear, there was no further effort to confirm the dates. If, however, there was a question about a date (especially what month people moved in or out when eligibility was close), then verification was requested. This generally took the form of rent receipts, utility bills, school records, or information from other family members. During each interview, the dates of residency were again asked and recorded. If these dates did not correspond to the dates given in the initial contact, the field supervisor sought clarification and made the final decision on eligibility.

B.2 Households in the Study

In Phase II two comparison households were included for each exposed one. Initially, interviews were only scheduled with members of exposed households who had indicated a willingness to participate in the Phase II study. Near the end of the summer, however, those who during the Phase I survey had expressed their unwillingness to take part in Phase II were contacted and invited again to participate in the research. At this point, a few households had changed their minds, and we did obtain additional agreements to participate. Difficulties in locating some comparison households, and the occurrence of refusals, resulted in the need for an expansion of the pool of households to serve as comparisons. In order to maintain a ratio of two

comparisons for each exposed household, the CDC provided a second set of comparison matches. The maximum sample size permitted under the contract was achieved before all of the households in this second set of comparisons could be contacted. Table B.1 shows the relation of Phase I contacts with households to the Phase II research.

Table B.1 Summary of Contacts by Household

	PHASE I					PHASE II				
	N	C	R	I	D	N	C	R	I	D
Verona -- Exposed *	68	61	3	1	3	69	49	13	4	3**
* 52 agreed to participate in Phase II										
** could not be contacted										
Springfield -- Exposed *	14	8	4	0	2	14	5	0	0	9**
* 5 agreed to participate in Phase II										
** could not be contacted										
Dowagiac -- Exposed *	16	16	0	0	0	17	12	5	0	0
* 14 agreed to participate in Phase II										
Battle Creek City -- Comparison *	94	59	25	0	10**					
* 179 completed interviews										
** 5 could not be contacted										
5 not needed due to sample size										
Townships -- Comparison *	44	25	9	0	10**					
* 108 completed interviews										
** 2 could not be contacted										
8 not needed due to sample size										
Ceresco -- Comparison *	22	14	4	0	4**					
* 49 completed interviews										
** not needed due to sample size										
Edwardsburg -- Comparison *	29	18	9	1	1**					
* 68 completed interviews										
** could not be contacted										
Niles/Barron Lake -- Comparison *	20	11	7	0	2**					
* 50 completed interviews										
** could not be contacted										

N = Sample size.

C = Completed interviews.

R = Refusals (may be ineligible).

I = Ineligible (because of dates of residency).

D = Dropped (could not contact, or not needed as sample size was achieved).

Table B.1 summarizes the contacts made with households for each of the three exposed areas and the five areas used for comparison. For each area, the total sample size is broken down into completed interviews, refusals, ineligible to participate (generally because they failed to meet the required dates of residency), and dropped cases (because they were impossible to contact or because the maximum sample size for the study had been achieved).

In Phase II, 100 exposed households were included in the study. Completed interviews were obtained from members of 66 (66%) of these households. Eighteen (18%) of the households refused to participate, four (4%) were found to be ineligible, and 12 (12%) were dropped because they could not be contacted. It must be emphasized that, as Table B.1 reveals, the 100 exposed households were composed of 71 households who had agreed in the Phase I interviews to participate in Phase II, along with prior refusals and two new cases. If completed interviews (N=66) are compared to the number of households who had previously agreed to participate in Phase II (N=71), a completion rate of 93% is obtained.

A total of 209 comparison households were selected for possible inclusion in the study. In fact, however, 17 of these households were never contacted because the maximum sample size for the study was reached before the research team had a chance to seek their participation. Of the 192 households who were contacted by the research team, completed interviews were obtained from members of 127 (66%) households, refusals were obtained from 54 (28%) households, one (<1%) household was found to be ineligible, and 10 (5%) were dropped because they could not be located.

As pointed out above, identical participation rates were achieved with exposed and comparison households. Completed interviews were obtained from two-thirds of the 100 exposed households and two-thirds of the 209 comparison households. Nonparticipants included actual refusals, households later determined to be ineligible, and households that had to be dropped (generally because they could not be located or contacted). Refusals, as expected, made up a larger proportion among comparison households. Unlike exposed households, many comparison households expressed the view that the water pollution was "not their problem" or that they just weren't involved. If ineligible and dropped cases are omitted, the differential rates of refusal by study area may be seen more clearly.

Table B.2 presents the interview refusal rate for households by study area. In this case the refusal rate is calculated as the ratio of the number of actual refusals to the number of refusals plus the number of completed interviews. Two general patterns appear to be present. As was expected, refusal rates were greater among comparison households than among exposed households, and refusal rates for both exposed and comparison households were greater in Cass County than in Calhoun County. This may well reflect the fact that water pollution problems in Calhoun County had received much more press coverage and probably had achieved greater salience in the public's view.

Participation in this research made many demands of study participants: lengthy interviews, release of medical records, participation in testing of blood and urine, collection of physical measurements, and permitting the testing of water supplies. Furthermore, the research took place, especially in Calhoun County, in a highly politicized environment that included a vocal citizen action group and actual and proposed lawsuits. Given this context and the level of these demands, the rate of completed interviews is quite satisfactory. Those who refused to participate were often vague about their reason for not wanting to take part. Of those who gave clearer statements, fear of the blood testing, the amount of time required, feelings of uninvolvedness, and a desire "not to be a guinea pig for the state" were among the reasons provided. Transportation problems were not a factor, as interviews were conducted at home, and the WMU project staff provided transportation to the clinic when necessary.

Table B.2 Household Interview Refusal Rate by Area -- Phase II.

Area	Completions	Refusals	Refusal per 100 completions
Verona -- Exposed	49	13	21
Springfield -- Exposed	5	0	0
Dowagiac -- Exposed	12	5	29
Battle Creek City -- Comparison	59	25	30
Townships -- Comparison	25	9	26
Ceresco -- Comparison	14	4	22
Edwardsburg -- Comparison	18	9	33
Niles/Barron Lake -- Comparison	11	7	39

These numbers and percentages are approximate since tracked households are not included, and each contacted address is counted as a household. Therefore, these results do not cross-reference with the CDC/MDPH report.

B.3 Tracking of Prior Households

One major aspect of the research involved the tracking of former residents of the exposed and comparison areas. Attempts were made to identify and locate previous residents of the exposed addresses and selected comparison addresses back to 1975. Once such persons were identified, located, and their eligibility established, they were asked to participate in the research. Interviews were arranged for such persons in their homes and provisions were made for collection of blood and urine samples and physical measurements. The tracked households interviewed were in areas in Michigan outside of Calhoun and Cass Counties and in eight states: Arizona, Arkansas, Colorado, Florida, Idaho, Indiana, North Carolina, and Texas.

A wide variety of techniques were used to identify and track former households. Among these were the use of city directories (for areas where they existed) for previous years, a Tracking Inquiry form that was used to solicit information on former residents from current residents, interviews with municipal clerks, a Battle Creek Inquirer article that included a list of 51 former residents of addresses included in the study whose current address was unknown, inquiries to relatives and former neighbors, and the checking of names against utility and telephone subscriber lists. Tracking activity was concluded in late November as the result of a joint decision by personnel from WMU, MDPH, and CDC when it was clear that the number of participants in the research had reached the maximum of 750 cases.

The addresses from which previous residents were to be tracked were provided by CDC. The actual number of prior households who had lived at each address was unknown and could only be determined through a complete enumeration of the households that had resided at the address throughout the full time period of interest. In some cases it was impossible to achieve a complete enumeration and the number of households identified represents a partial listing of the households of interest. Of the 66 addresses that were provided for tracking, all households who had lived at an address during the period of interest were identified for 50 (76%) of the addresses. In the remaining cases, information was generally available for some of the years during the period of interest. Table B.3 provides a summary of tracking activity for the 66 tracked addresses by study area.

Table B.3 Tracking activities for former households of the study areas.

Area	Number of new addresses to be tracked	Number of addresses identified	Number of addresses interviewed	Reason for no interview			
				CNL	D	R	I
Verona	14	28	13	6	7	1	1
Springfield	1	0					
Dowagiac	4	6	4	1	1		
Battle Creek City Townships	32	60	23	18	4	3	12
Ceresco	4	5	1	1	2	1	
Edwardsburg	4	6	3	3			
Niles/Barron Lake	1	0					

CNL = Could not locate.

D = Deceased, next of kin not found or unable to give information.

R = Refusal.

I = Ineligible.

It is not known whether households not located were eligible for participation in Phase II. unknown. As stated for Table B.2 these numbers are approximate since tracked households are not included and each contacted address is counted as a household.

B.4 Interviews

The interview schedule used in the study was developed in cooperation with MDPH and the CDC. The 24 page instrument collected data on demographic variables; previous residency; water use patterns; residential exposure to solvents, glues or resins, paints, pesticides, chemical treatments for yards, insecticides, and weed killers; occupational history, occupational exposure to chemicals; use of alcohol and tobacco; use of prescription and "over the counter" medications; and history of contacts with physicians for a wide range of problems including various skin conditions, liver problems, blood and vascular disorders, urinary or renal conditions, nervous system problems, metabolic conditions, allergic conditions, cancer, and any other diseases. Information was also collected on the presence of symptoms such as headaches, fatigue, nausea, malaise, nervousness, irritability, and sleepiness; and on the family history of cancer, liver problems, yellow jaundice, and diabetes. For female respondents, an additional series of questions obtained information on prior pregnancies and births and on current use of oral contraceptives. Additions to the interview schedule were made at the request of CDC on July 30 (a series of questions asking if medical conditions had been seen by a physician before 1970) and on August 15 (two questions about consumption of fish).

The interviews were conducted by 23 interviewers selected from among applicants by the Principal Investigator and Field Supervisor. In addition, the Tracking Coordinator and the Field Supervisor conducted a small number of out-of-state interviews. Interviewers were typically graduate students or graduates, and several held advanced degrees. Many of the interviewers had prior experience in survey research on earlier research projects conducted by the Kercher Center for Social Research or similar employers.

Interviewers on this project participated in two half-day training sessions prior to conducting any interviews. These sessions covered a review of basic interviewing procedures, an overview of the research project, an explanation of how research subjects were being scheduled for interview, an explanation of the informed consent and medical release forms, a detailed review of the interview schedule, and practice in administering the schedule. Each interviewer was provided with a letter of introduction from the Principal Investigator and with a supply of the Principal Investigator's business cards to be distributed so that any study participant with additional concerns or questions could call the study director. The interviewers were continuously monitored by the Field Supervisor throughout the period of data collection. Interview schedules were returned to the Field Supervisor for review the day after the interview.

Potential participants in the study were initially contacted through a letter from the Principal Investigator. This letter set out the purpose of the study, and explained that the Kercher Center for Social Research was conducting the study for the Michigan Department of Public Health. The letter also explained the general content of the questions to be asked in the interview and informed that a member of the Center staff would be contacting them to arrange for an interview and to answer any questions they might have. Copies of the Statement of Informed Consent and Patient Information Release Authorization were also enclosed in the letter in order to permit the person adequate time for review. The letter included a brief overview of these forms and an explanation that the participant would be asked to sign the forms at the time of the interview. The letter explained that upon completion of the visit to the laboratory site, each individual who participated in the research would receive \$50 to compensate for the time involved in the research. Finally, there was an invitation to call if there were any questions about the study.

The letters were followed by a phone call from a member of the project staff who attempted to answer any questions the potential study participant might have, screened for eligibility, and scheduled interview appointments for all eligible household members. These interviews were scheduled at the convenience of the persons to be interviewed. Research participants who were 10 years of age or younger had the interview completed by a parent. Youths over 10 were interviewed alone or with parents, as the youth wished. Interviewers were instructed to ensure that older youths have privacy for answering questions where accurate responses might be inhibited by parental presence. Virtually all of the interviews were conducted in the homes of the study participants. The only exceptions to this were one interview conducted at the subject's office, one interview conducted in a subject's hospital room, two interviews conducted at the project office at Western Michigan University, and two out-of-state interviews were conducted by telephone. Interviews with participants in the research began on July 11, 1985 and were completed during December of 1985.

Generally, of course, study participants personally provided all the responses to the survey instrument. Of the total of 749 completed interviews, 601 (80%) were personally answered by the actual research participant. In the case of all children ten and under and some older children, the interview schedule was completed by a parent (generally the mother). A total of 121 children had their interview schedules completed by a parent. In addition, 15 schedules for deceased persons were completed by a next of kin. In a few other cases, the spouse or other family member answered the survey questions for a participant who was incapacitated. Table B.4 shows the relationship of those who actually answered the survey items to the research subject.

Over 90 percent of the interviews were conducted in the exposed and comparison group areas. Thirteen interviews were conducted with people who still resided within Michigan but outside the major study areas. A total of 31 interviews (two by phone) were conducted with respondents who now reside in other states.

The \$50 participant incentive fee for adult study participants was paid upon completion of the blood and urine sampling/physical measurements portion of the research. The incentive fees for children were paid upon completion of the laboratory visit by the last adult member of a household. For participants who completed the interview and provided access to medical records but declined to provide blood and urine samples, a \$15 incentive fee was used. For convenience of the study participants, the WMU project office established local accounts in Battle Creek and Dowagiac, and the incentive fees were provided through checks written on these accounts. The checks were given to study participants at the lab sites upon completion of their visits.

Table B.4 Relationship of Respondent to Study Subject.

Relationship	Number of Cases	Percent
Self	601	80.2
Spouse	12	1.6
Parent	121	16.2
Child	6	0.8
Other	9	1.2

Table B.5 summarizes the distribution of the 749 completed interviews. A total of 251 interviews were completed in the three exposed areas and a total of 498 interviews were completed in the five comparison areas.

Table B.5 Completed Interviews (total = 749) by Area.

Exposed areas		Comparison areas	
Verona	181	Townships	87
Springfield	16	Ceresco	43
Dowagiac	54	Battle Creek City	267
		Edwardsburg	67
		Niles/Barron Lake	34
Total	251		498

Once the completed interviews were collected, they were checked in by the field director who also examined them for completeness and accuracy. They were then given to the coding supervisor for assignment to coders. Since the data to be transferred to mark-sense sheets were both voluminous and complex, the coding supervisor performed a complete verification check on the initial assignments of each coder, and errors were immediately corrected. Following the initial assignments, the coding supervisor performed random checks on all cases to insure that the level of accuracy was being maintained. Once all the data were coded and loaded onto the computer system, a frequency distribution was charted for all variables. This allowed the identification of any out of range codes that had escaped the verification process during coding. Whenever such errors were discovered, the case was rechecked record by record.

B.5 Project Staff

Principal Investigator: James C. Petersen, PhD
Senior Data Management Specialist: Thomas L. Van Valey, PhD
Field Supervisor: Janet Van Valey, MA
Tracking Coordinator: Tudy Boldin, BA
Secretary: Lois Carl

Interviewers: Twenty-three interviewers were hired to conduct interviews with the study participants. A few out-of-state interviews were conducted by members of the project staff. Most interviewers were graduate students or recent university graduates with prior research experience.

B.6 Blood Sampling and Laboratory Work

Borgess Medical Center, our subcontractor for the collection and analysis of blood and urine samples and physical measurements, established field collection sites at Leila Hospital in Battle Creek and Lee Memorial Hospital in Dowagiac. Over 90% of the 692 study participants who completed the lab testing/physical measurement portion of the study were seen at these two sites. The remaining participants were either seen at Borgess Medical Center or were participants who had moved outside the study area. In these cases, Borgess Medical Center selected an appropriate location for local testing, sent detailed instructions and return mailers, and received the samples for testing by express mail. The WMU project staff established the appointment schedules for research subjects and prepared the checks used to compensate research subjects for the time involved in the study.

Of the 749 research participants for whom completed interviews were obtained, 692 completed the blood-urine sampling/physical measurements portion of the research (three without blood samples). Of the 57 who did not complete this aspect of the research, 24 were under five years of age, 15 were deceased, 17 refused this part of the study, and one was not scheduled.

As part of the quality control program, additional blood was obtained from every tenth research subject to permit the establishment of duplicate blood samples that were assigned unique identification numbers. The numbers for these quality control cases were assigned by the WMU research staff using procedures that made these numbers indistinguishable from the normal identification numbers. A revised informed consent form, that permitted the withdrawal of a larger volume of blood was obtained from every tenth research participant.

The first field collection site was held on August 19, 1985. The final local samples were obtained during December, and the final mail-in samples were received during January, 1986. Table B.6 shows the distribution of cases with laboratory data by location at which the blood/urine samples were collected.

Table B.6 Specimen Collection/Physical Measurements by Location.

Battle Creek -- Leila Hospital	512
Dowagiac -- Lee Memorial Hospital	118
Kalamazoo -- Borgess Medical Center	21
Out of Area -- Mail-Ins	41
Total	<u>692</u>

B.7 Abstracting of Physician and Hospital Records

The abstracting of the previous physician and hospital records of participants was conducted by experienced medical records professionals at Borgess Medical Center. Sharlene K. Dolman, RRA, served as director of this project. Other Borgess personnel who abstracted records included Barbara Swanson, RRA; Mary Kenney, ART; and Ceil McDermott, RRA.

The names and addresses of physicians and hospitals previously visited by study participants were obtained from the responses provided during the participant interviews. In addition, the names of all study participants were provided to area hospitals with a request to search their records to determine if any of those listed had been admitted to the hospital. The hospitals that complied with the request for a records search were: Lakeview General Hospital, Battle Creek Adventist Hospital, Veterans Administration Hospital (all three in Battle Creek), Pawating (Niles), Lake View Community (Paw Paw), Saint Joseph's Medical Center (South Bend, Indiana), Michiana Community (South Bend, Indiana), and Elkhart General (Elkhart, Indiana).

When possible, Borgess personnel made personal visits to physician offices or hospitals to abstract records. When the distance made this impossible, copies of full records were obtained by mail and abstracted at Borgess Medical Center. Separate forms listing the diagnoses made at each location were prepared for each study participant. At the conclusion of the abstracting, a summary abstract was prepared for each participant which included diagnoses, ICD9-CM codes, the date of each diagnosis, and the location where it was made (physician office or hospital).

In cases where a reduction of the number of diagnoses was possible, in cases of "ill-defined disorders," and in instances where the date of first diagnosis was unclear, the abstracts were reviewed by the chair of a medical panel, Dr. Jack Hunt, a board certified family practitioner. With cases which he felt were not clear cut, two additional board certified family practitioners, Dr. George Lode and Dr. William Bateman, independently examined the abstracts. Dr. Hunt made the final determination in cases where there was disagreement among the panel members. Medical abstracts were obtained from records at 75 hospitals and 276 physician offices. Thirty physicians identified by participants were not contacted because they had retired or could not be located. Nineteen others (involving 34 study participants) refused to cooperate in the study. A total of 2,073 records were abstracted during the project.

B.8 Costs of data collection

The Michigan Department of Public Health has received official cost breakdowns, by contract category, of all funds spent on the research project. Rather than repeating these reports, which were prepared by the University's Grants and Contracts Office, a summary was constructed of the costs related to the various forms of data collection in the WMU part of the project. In compiling these figures, no attempt was made to distribute the cost over factors such as supervisory personnel, clerical personnel, general supplies and services, coding of data, data entry, and computer processing. Since WMU study personnel generally worked with all forms of data collection and the computer files that were produced encompassed several forms of data, it would be difficult to accurately assign costs to a single type of data collection. Further, the dollar amounts refer to direct costs, except with the Borgess subcontract where the Borgess indirect costs were included. WMU charged an indirect rate of 20% on all the project costs, except those related to the Borgess subcontract.

B.8.1 Interviews

The direct cost of interviewing and travel in the local study area equalled \$24,960. A total of \$7,395 was spent in direct cost for interviewer time and travel outside the local study area. In addition, a substantial amount of WMU staff time was spent in scheduling interviews and interviewers.

B.8.2 Tracking activity.

A total of \$17,191 was spent in direct costs for the tracking coordinator, telephone, postage, and travel. In addition, some tracking questions were asked by interviewers at the conclusion of their regular interviews. There was a moderate amount of additional support provided by the WMU research staff.

B.8.3 Abstracting of physician/hospital records

The portion of the Borgess Medical Center subcontract that was related to the abstracting of records equalled \$89,700. In addition, some support for this activity was provided by the WMU research staff. This activity was largely confined to xeroxing of portions of the interview forms, obtaining updated medical release forms, and supervision.

B.8.4 Operation of field clinics

The operation of the two clinics for the collection of blood-urine samples and physical measurements, and the arrangements for the collection of out-of-area samples accounted for \$11,344 of the Borgess Medical Center subcontract. In addition, major support was provided by the WMU staff which did the scheduling of laboratory appointments, provided some transportation to study participants when needed, and supervised the efforts.

B.8.5 Laboratory testing

The total costs attributed to Borgess Medical Center and its subcontractor for the testing of blood and urine samples equalled \$105,548. WMU staff involvement was largely confined to assisting MDPH in providing laboratory results to participants and supervision.

B.8.6 Participant incentive fees

A total of \$36,274 was expended in the form of participant incentive fees and related costs that were incurred in the form of fees and charges for checks in the Battle Creek and Dowagiac bank accounts. In addition, a significant amount of WMU staff time was used to prepare and distribute checks and to maintain the documents used to record receipt of payment.

Appendix C

Estimation of Individual Exposure to Volatile Organic Chemicals (VOCs) for Residents with Contaminated Wells in Neighborhoods Adjacent to the Verona Well Field, Battle Creek, Michigan

Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia

and

Donald L. Phillips, PhD
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia

A modified and expanded (for the purpose of reporting the
Battle Creek Health Study) version of a scientific paper entitled
Estimation of the Time Component in the Movement of
Chemicals in Contaminated Groundwater
Environmental Health Perspectives 74: 211-221, 1987

C.1 Introduction

In conducting an epidemiologic analysis, exposure is usually considered to be constant over time, unless data to the contrary are available. Little attention has been given to the time factor in environmental studies. Yet, in risk assessment and epidemiologic studies, the time factor is important in a number of ways:

- * A disease diagnosed before the date exposure began cannot be attributed to the exposure.
- * The date that exposure started determines whether the observation time was long enough for a disease to be attributable to the exposure of interest.
- * The date that exposure stopped marks the beginning of the exposure-free observation period, in which the dose-response curve for chemicals that are readily excreted may be different from that when exposure was still ongoing (1).
- * The dates when exposure started and ended, and the changes in the exposure level between these two points in time (exposure as a function of time), have to be known for the estimation of the duration and total amount of exposure. Both are important for a proper epidemiologic analysis of the association between exposure and disease.
- * The age at which exposure started and ceased is necessary information in estimating the risk of cancer through mathematical models for a less-than-lifetime exposure (2).

Clearly, in epidemiology, neglecting the time factor is likely to result in misclassification with regard to the exposure status and the disease status of individuals, and can, thus, lead to decreased statistical accuracy of studies and to erroneous inferences.

In the case of contaminated groundwater, once a chemical spill has occurred, a certain period of time (dT_1) is needed for a chemical to reach a distant well (Figure C.1). The magnitude of dT_1 is determined by many factors, such as depth of the water table, soil structure, direction and velocity of the groundwater flow, the distance of the well from the main axis of the flow and from the spill site, withdrawal rate and amount of the well, soil affinity and other properties of the contaminant, etc. Once the compound reaches a well at the time T_1 (the time that a compound reaches the detection limit), a period of time dT_2 is needed to reach a sample concentration C_s at sampling time T_s (Figure C.1). The estimation of T_1 , dT_1 , and dT_2 is retrospective in situations of groundwater contamination. Estimating these values is important in tracing the source of the pollution, in seeking recovery for cleanup costs or material damage, and in determining possible adverse health effects. Retrospective estimation of T_1 and $C=f(T)$, a mathematical expression of the changes in the chemical concentration C as a function of the time T_1 , is necessary when the potential health effect of chemicals in water is to be studied. In the typical scenario of an incidentally detected pollution of wells, however, existing information will probably be limited to the results of recent tests of water samples, since routine monitoring of public water supplies for a wide variety of toxic chemicals began only a few years ago.

In theory, predictive models can be used retrospectively. Several models have been advanced for mathematically describing the transport of chemicals in groundwater. However, all of these require extensive data on hydro-geologic parameters, which were not available for the situation described in this paper, and they do not exist in virtually all scenarios of incidentally detected contamination of groundwater. Moreover, as described, such models are inherently incapable in discriminating between individual wells in a crowded well field (See Discussion). This paper describes the development and use of a method for retrospectively estimating T_1 , $C=f(T)$, and the individual exposure status.

The method is based on the hypothesis that monitoring data for city wells can be used for estimating $C=f(T)$ for city wells. This equation can then be applied to neighboring residential wells for the retrospective estimation of T_1 and the individual exposure status as depicted in Figure C.1.

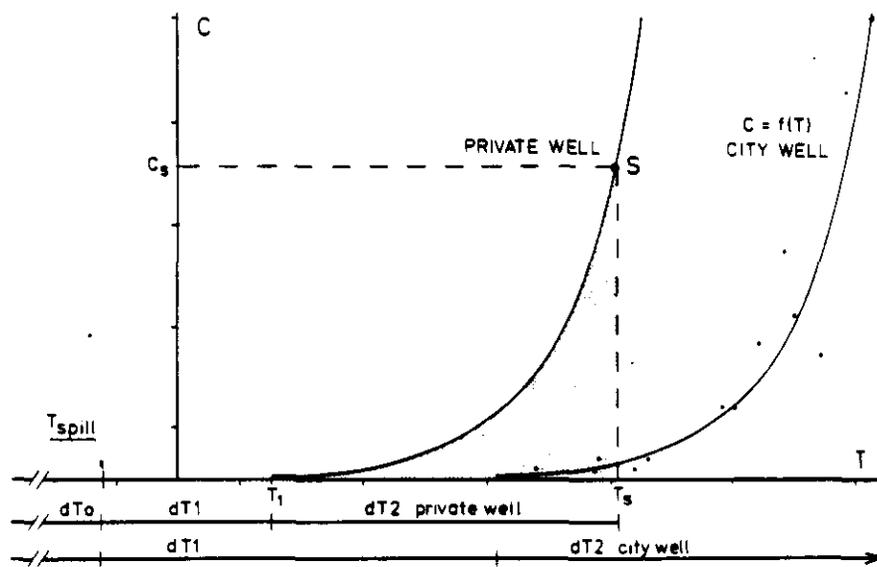


Figure C.1 Application of $C=f(T)$ from city wells to sample concentration C_s at time of sampling T_s of a private well. The dots on the right side of the figure represent actual values of a time series for trichloroethane (TCA) in city well #28. The right curve represents $C=f(T)$ for the city well and TCA. By shifting the curve to the left through sampling point S, time T_1 (when TCA reached a concentration of 1 part per billion) can be calculated for the private well. The time lapses dT_0 , dT_1 and dT_2 are determined by the time the spill began (T_{spill}), T_1 , and the time of sampling T_s .

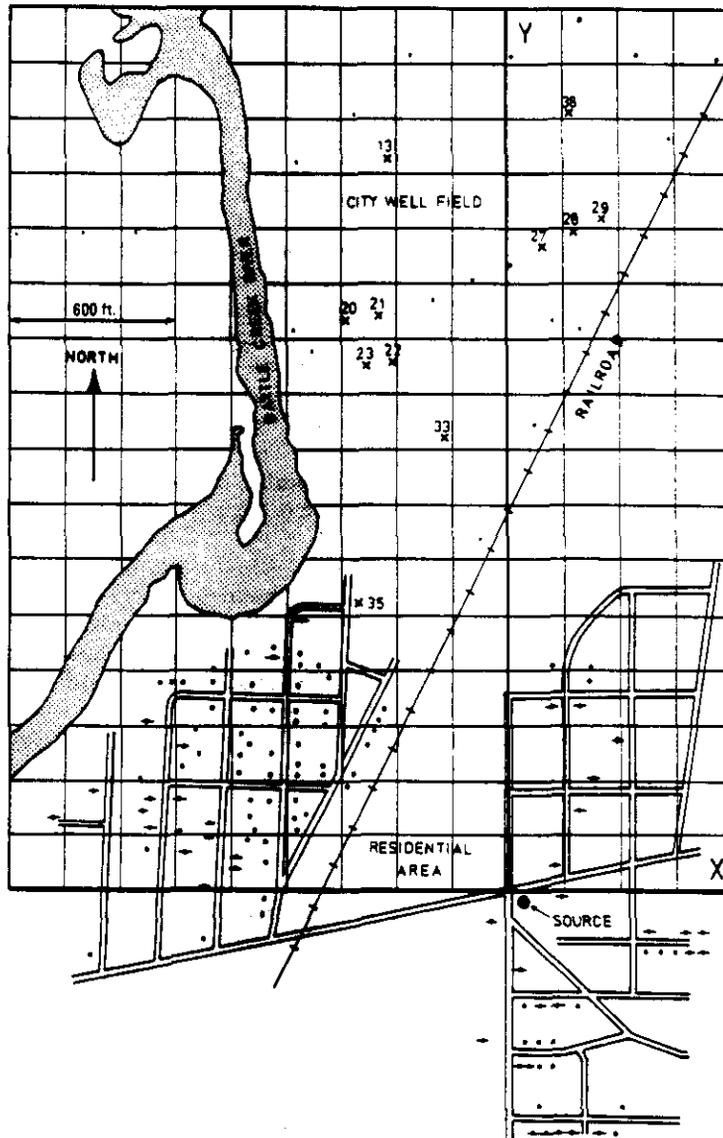


Figure C.2 Area map of the Verona city well field and the adjacent Verona Park private wells. The X-Y grid system is used to express the distance of the wells from the source of the spill. The city wells used for developing $C=f(T)$ are numbered. The (--) marked private wells had no VOCs at the time of sampling.

C.2 Materials and Methods

In September 1981, an unscheduled survey led to the discovery of VOCs in the municipal drinking water of Battle Creek, Michigan. A routine monitoring program for city wells was set up, and neighboring private wells were tested (only one time in the period 1981 - 1983). Contamination of a number of wells was confirmed. There were no historical data on the presence of VOCs in city or private wells. The source was found to be a distributor of industrial solvents, who was located up-gradient of the affected wells (3). The major source of the spill appeared to be leaking underground storage tanks. Soil samples taken at the site showed extremely high concentrations, and VOCs were found floating on top of the water table. Thus, the contamination of the aquifer appeared to be from continuous spilling rather than a one time spill. Figure C.2 depicts a map view, constructed from an aerial photograph, and from street maps supplied by the city engineer.

Groundwater flows in three aquifers in the Verona area. From the land surface down to the aquifers are sand and gravel units within unconsolidated glacial deposits, and the upper and lower sandstones of the Marshall Formation. Horizontal hydraulic conductivities of the aquifers are about 100, 150, and 550 feet per day for the sand and gravel, upper sandstone, and lower sandstone respectively. Secondary permeability of the sandstones, due to fracturing, is the cause of their relatively high hydraulic conductivities. Most private well water is drawn from the sand and gravel or the upper sandstone aquifers. Municipal water is drawn from the upper and lower sandstone aquifers. Although the location of the river west of the wells would lead one to expect groundwater to flow from the point source in a west to northwesterly direction, the actual flow was found to be in a north to northwesterly direction. This can be explained by the heavy water withdrawal in the city well field. The velocity of the flow has been estimated at 1 to 4 feet a day. The above hydrogeologic information is abstracted from a study by Grannemann and Twenter, conducted when the contamination of the city well field was detected (4).

Estimation of T_1 requires determination of $C=f(T)$ and, thus, repeated testing of water from the same well over time. Such time series, covering a period from September 1981 to the present, were available for 30 city wells, but not for the 116 private wells tested. Data on well descriptors and water quality were provided by the Michigan Department of Public Health and the Department of Natural Resources. The VOCs found most frequently, their abbreviations as used in the text, and the ranges of concentrations are depicted in Table 1 (Chapter 1). The following sections describe the steps in determining $C=f(T)$ for city wells. This equation is then applied to individual private wells to estimate T_1 retrospectively.

C.2.1 Selection of wells

Minor remedial changes in the management of the city well field started in 1982, and reached major proportions in 1984 when newly constructed city wells north of the well field came into operation. The new city wells caused, as intended, a sharp drop in VOC-concentrations in the existing wells, and the flow pattern in the aquifer was drastically altered (5). Since declines in the level of concentration were the results of the change in the management of the city well field, a change not occurring in residential areas, it was presumed that residential wells continued to be affected by increasing levels at least until mid 1984. Accordingly, the city wells selected for developing a VOC-specific $C=f(T)$ were limited to wells showing increasing VOC concentrations over one year or more during the interval 1981 to mid 1983. Not all wells were polluted and if they were, contamination did not necessarily involve all VOCs. In particular, 1,2-dichloroethane (12DCA) was rarely found in city wells, but it was common in private wells. Table C.1 presents a summary of the VOC-levels in city and private wells.

Table C.1 VOC levels in the groundwater of the Verona city well field and the adjacent private well area.

VOCs**	Number of contaminated private wells (total N = 90) *					Number of contaminated city wells (total N = 27) *						
	Sample concentration (ppb)	0	1-4	5-24	25-99	100+	Maximum concentration (ppb)	0	1-4	5-24	25-99	100+
11DCA	38	17	17	16	2	7	8	7	5	0		
12DCA	42	16	12	16	4	23	3	1	0	0		
DCE	56	20	11	3	0	14	7	6	0	0		
CIS	24	14	14	7	31	11	2	9	3	2		
PCE	53	11	8	11	7	10	7	3	6	1		
TCA	46	23	16	4	1	7	6	7	5	2		
TCE	47	8	8	21	6	13	12	0	2	0		

* In addition to the wells cited, 3 city wells were free of VOCs during the study period, and 48 private wells tested negatively for VOCs in the one each sample taken in the same period.

Of the 30 city wells, 11 were either not polluted, tested positive a few times only, or became contaminated as late as 1984. In these 11 wells, VOC levels ranged from 1 to 11 ppb; in none of these wells were more than one or two VOCs present. Of the remaining 19 wells, eight showed low VOC-levels with single sample peaks, or the concentration versus time curve (CT-curve) did not show an increasing trend. Eventually, 11 city wells provided a total of 27 time series suitable for modeling. These wells are depicted in Figure C.2 (numbered wells).

C.2.2 Variables for developing a time-concentration model

A grid pattern with a central X-Y cross-axes system was superimposed upon the map of the study area, with a north-south Y-axis. The intersection point of the axes is located just northwest of the spill site (Figure C.2). The coordinates for each well were expressed as the number of grid intervals from the X and Y intersection point. The rationale for the X-Y coordinates was the hypothesis that, although the movement of the plume of contaminants would follow the main direction of the groundwater flow (Y-axis), microscale deviations were likely to occur from diffusion or local differences in the soil structure or flow characteristics. Such deviations are projected as components of X and Y. In addition to X and Y, the depth of the well was important, since a downward component of the groundwater flow and a vertical gradient in VOC levels were found in hydrologic studies (5).

When a time series of a particular city well showed an increasing trend followed by a decline in 1984 (due to changes in aquifer management), the descending part of the series was deleted before statistical processing. Thus, no time series extended beyond April 1984. The time series used for developing $C=f(T)$ was arranged as a line-listing, in which each water sample was a single observation with the following information:

- The well number (for city wells) or street address (for private wells);
- The X and Y coordinates, and a number for the quadrant of the XY system;
- The month and year of sampling, from which the variable T_s was calculated as the number of months since January 1, 1970, an arbitrarily chosen date;
- The sample concentration C_s of each chemical in ppb (parts per billion);
- The well depth D in feet (midpoint of the screened or uncased portion);
- The well pump capacity G in gallons per minute.

C.2.3 Development of the time-concentration model

The statistical models tested were based on linear regression of C on T as of a common starting date, arbitrarily set at January 1, 1970. This date is one and a half years after the solvent company started its operations. This time lapse is equal to dT_0 plus dT_1 (Fig. C.1), and was considered a minimum length of time required for the spilling to occur (since the solvent company started its operations), and for the spilled VOCs to reach the wells at the detection limit. Most city wells showed VOC-positive water samples later than September 1981, implying that in the respective time series one or more leading samples were free of VOCs ($C_s = 0$). In such cases, all but the last leading zeros were censored. The last leading zero, and zeros occurring after at least one positive sample, were converted to an arbitrary $C_s = 0.5$ ppb to allow the transformation of C into $\log C$ (the natural logarithm of C); and because 0.5 is closer to the detection limit (between 0.2 and 1 ppb). Univariable regression (the only regressor variable being T) was applied to determine the model that best fitted the individual time series. Multivariable models were employed for data pooled from all selected wells, using stepwise backward elimination of insignificant variables with a threshold of $p = 0.1$ (6). The following basic models were tested:

univariable models

$$C = a + bT$$

$$\log C = a + bT$$

multivariable models

$$C = a + b_1T + b_2X + b_3Y + b_4D + b_5G$$

$$\log C = a + b_1T + b_2X + b_3Y + b_4D + b_5G$$

In the above models, "a" is the intercept and "b" the regression coefficient. In this paper, b_1 (the regression coefficient of T) will also be referred to as the slope of the regression, and the models will be referred to as $C = T$ and $\log C = T$ models, respectively.

C.2.4 Application of models for estimating T_1

Let C_s and T_s denote the concentration and time of a water sample from a private well, and T_{low} the time in the past when a chemical reached the lower concentration C_{low} of interest. T_{low} can then be estimated by selecting the best model from the stepwise regression for a given chemical, say, a $C = T$ model comprising T, X , and Y for TCA, and applying it to the target private well for both C_s and C_{low} . Then, subtract and solve for T_{low} :

$$C_s = a + b_1T_s + b_2X + b_3Y$$

$$C_{low} = a + b_1T_{low} + b_2X + b_3Y$$

Subtraction of the second equation from the first results in:

$$T_{low} = (C_{low} + b_1T_s - C_s) / b_1 \quad (\text{Eq 1})$$

The same derivation can be used for a log C = T model to show that:

$$T_{low} = (\log C_{low} + b_1 T_s - \log C_s) / b_1 \quad (\text{Eq 2})$$

Equations 1 and 2 show that all variables but T drop out because of their constancy over time. It is therefore not necessary to know a , b_2 , and b_3 for the individual well and chemical compound of interest. Since C_{low} represents any concentration for which the associated T_{low} has to be estimated, and thus also $C_{low} = 1$ ppb, T_1 can be estimated by replacing C_{low} with 1, which results in the following derived equations:

$$T_1 = (1 + b_1 T_s - C_s) / b_1 \quad \text{for} \quad C = T \text{ models} \quad (\text{Eq 3})$$

$$T_1 = T_s - (\log C_s / b_1) \quad \text{for} \quad \log C = T \text{ models} \quad (\text{Eq 4})$$

C.2.5 Evaluation of the validity of the models

The following procedures were used to select the best model for each chemical:

- 1) Judge for each single city well and VOC how well univariable $C = T$ and $\log C = T$ models fit the data from which they were derived, utilizing the r -value and the distribution of residuals. The results of the univariable analyses are shown in Table C.2 for each of the 27 time-series.
- 2) Pool the well data and judge how multivariable regression models fit the pooled city well data from which they were derived, utilizing the r^2 value and the distribution of residuals. With observed T_{low} and C_{low} equal to the time and concentration of the first positive sample in a time series, project for each subsequent sample of the time series an expected T_{low} using Equation 2. The magnitude and the distribution of the differences between predicted and observed T_{low} are good indicators of the validity and precision of a model for estimating T_{low} .
- 3) Repeat the previous step by comparing the observed and estimated T_{low} for the private wells that were sampled twice before 1984. The results of this comparison are shown in Table C.3.
- 4) Compute T_1 for the other private wells. As these wells lack a reference value of a second sample, a relatively narrow (a few years) range of T_1 estimates for a group of neighboring wells should serve as an alternative evaluation criterion for the validity of the model. The results of this evaluation is shown for 11DCA and CIS in Figure C.3.
- 5) Repeat steps 2 through 4, now using weighted regression models, in which the data from an individual well are weighted by $1/SE^2$ of that well, where SE is the standard error of the slope of the univariable regression.

Table C.2 Wells, VOC, VOC-specific slopes (b_1), and correlation coefficients (r^2) based on univariable regression of C (oncentration) on T (ime).

City Well No.	VOC	$C = T$ model		$\log C = T$ model	
		b_1	r^2	b_1	r^2
13	CIS	1.860	0.66	0.375	0.95
20	CIS	0.758	0.55	0.186	0.67
21	CIS	0.982	0.72	0.148	0.77
22	11DCA	0.272	0.80	0.118	0.80
	CIS	3.018	0.90	0.174	0.65
	TCE	0.631	0.73	0.194	0.65
23	CIS	1.965	0.59	0.187	0.49
27	11DCA	0.839	0.71	0.053	0.76
	DCE	0.259	0.46	0.067	0.41
	TCA	5.013	0.78	0.110	0.87
	PCE	2.931	0.78	0.094	0.86
28	11DCA	0.606	0.52	0.095	0.70
	DCE	0.172	0.37	0.076	0.50
	TCA	4.880	0.65	0.166	0.95
	PCE	2.858	0.59	0.152	0.93
29	11DCA	0.337	0.64	0.099	0.57
	TCA	2.590	0.85	0.177	0.85
	PCE	1.533	0.88	0.163	0.85
33	11DCA	0.500	0.69	0.095	0.76
	CIS	5.539	0.91	0.167	0.94
	TCE	1.035	0.91	0.131	0.92
35	12DCA	0.681	0.51	0.105	0.36
	CIS	5.286	0.30	0.032	0.23
38	11DCA	1.031	0.80	0.200	0.86
	DCE	0.261	0.71	0.163	0.74
	TCA	0.824	0.67	0.150	0.76
	PCE	1.706	0.88	0.270	0.93

The number of samples per series varied from 9 to 19 and the length of sampling from 10 to 19 months. For well location, see Figure C.2.

C.3 Results

The following paragraphs refer to the results of applying the five evaluation procedures for evaluating the model fit, described above.

- 1) Of the 27 VOC- and well-specific time series, most showed curved scatter plots of C versus T , resulting in 17 series with higher r^2 values for $\log C = T$ models, two series with equal r^2 for $C = T$ and $\log C = T$ models, and eight series with higher r^2 for the $C = T$ model (Table C.2). Figure C.1 depicts a typical $\log C = T$ fitting scatter plot, representing the time series of city well # 28 for TCA. The $C = T$ models showed a tendency towards underestimating C at both ends of the curve, especially at the upper end, and towards overestimating C in the middle part of the curve. The $\log C = T$ models tended to slightly overestimate C at the lower end of the curve.
- 2) In general, the results of the unweighted multivariable regression models were again in favor of $\log C = T$ models. $\log C = T$ models had a higher r^2 for TCA, PCE, and 11DCA, whereas DCE, CIS, and TCE had a higher r^2 with $C = T$ models. No multivariable regression model was developed for 12DCA, a chemical found only occasionally, and at very low levels, in city wells. The maximum concentration was 10 ppb. With regard to the difference between observed and estimated T_{low} , $C = T$ models showed increasingly large differences with increasing C with large variations (mean difference of +2.6 months, standard deviation or S.D. = 24.5). In contrast, the differences resulting from $\log C = T$ models scattered, within a narrow range, around a mean of zero months along the time-axis with S.D. = 6.5. In 56% of the predictions, the estimated T_{low} was closer to the observed T_{low} if a $\log C = T$ model was used, compared with 32% in case of using a $C = T$ model. No difference between the two models was seen in 12% of the T_{low} estimates.
- 3) As another test of their accuracy, the models were applied to the two private wells which, coincidentally, were sampled twice before 1984. Table C.3 shows that in 10 of the 14 pairs of observed and estimated T_{low} estimates derived from the $\log C = T$ model were the closest to the observed value. The range of differences between estimated and observed T_{low} is by far the widest for estimates derived from $C = T$ models. In one case the $C = T$ model yielded a T_{low} estimate pointing to a time still to come (in 1997).
- 4) Table C.4 depicts T_1 estimates for private wells derived from unweighted multivariable models. $C = T$ models yielded the widest range of estimates (134 years) and many large negative values, that is, T_1 estimates far earlier than 1970, in extreme cases many decades to a century, even before the VOCs came in commercial production. T_1 estimates for neighboring wells differed from many years to decades with no consistent pattern. In contrast, T_1 estimates from $\log C = T$ models ranged from 1977 to 1983. Figure C.3 maps the location of private wells with their T_1 estimate (or observed if the sample concentration was 1 ppb) in periods of 1 year, for CIS and 11DCA using $\log C = T$ models. The close correlation of each well's T_1 with its topographical location is reflected in the clumping of the earliest T_1 estimates in an area closest to the spill site. From here, protrusions with later T_1 estimates spread out, pointing mostly in a northwesterly direction, in agreement with the main direction of the groundwater flow.
- 5) When analyzing pooled data, weighted regression models are conceptually preferred over unweighted models. Indeed, r^2 values and the accuracy in estimating T_{low} improved for $\log C = T$ models, although not for $C = T$ models. However, weighted models resulted in larger differences between observed and predicted T_{low} for both city wells and the two private wells in Table C.3. Apparently, giving more weight in the pooled data to wells with smaller standard errors of the slope does not guarantee more accurate estimates.

Table C.3 Observed and estimated T_{low} of private wells sampled twice before 1984. Because the first sample of well 260 has a C_{low} of 0 ppb of TCA, T_1 was estimated rather than T_{low} . The observed reference for this estimate is between T_{low} and T_{high} .

	11DCA	12DCA	DCE	CIS	PCE	TCA	TCE
Residential well 260							
Observed $C_{low} - C_{high}$	36-58	119-284	1-16	1100-674	16-42	0-12	37-46
Observed $T_{low} - T_{high}$	145-155	145-155	145-155	145-155	145-155	145-155	145-155
Estimated $T_{low} (C = T)$	116	60	85	306	144	$T_1 = 151$	145
Estimated $T_{low} (\log C = T)$	150	151	142	158	149	$T_1 = 139$	153
Residential well 620							
Observed $C_{low} - C_{high}$	13-16	17-24	3-4	549-488	55-77	2-4	60-68
Observed $T_{low} - T_{high}$	145-147	145-147	145-147	145-147	145-147	145-147	145-147
Estimated $T_{low} (C = T)$	142	137	142	169	137	146	138
Estimated $T_{low} (\log C = T)$	145	144	143	148	145	142	146

From the above, it was concluded that unweighted $\log C = T$ models yielded the best results when retrospectively estimating T_1 and $C = f(T)$. The widest confidence range for T_1 estimates was for chemicals with the smallest number of time series suitable for modeling. Confidence limits around a regression estimate would set limits around C_s and not T_1 , rendering conventional methods for estimating these limits not applicable. Since calculation of both C_s and T_1 are based on b_1 , the estimated regression coefficient for T , a sort of "95% confidence range" was established by using b_1 plus or minus 1.96 times its standard error. Ignoring 12DCA, the standard error (SE) of b_1 for $\log C = T$ models was not more than 6 to 15% of b_1 , resulting in a "confidence range" of a few months to a year, with an occasional outlier of two years. This is quite satisfactory, in view of the period covered (1970-1984) and the limited data available. The models for 12DCA were derived from one time series only (well #35), with a very low level of pollution and a monitoring period of only one year, which resulted in a SE of the regression slope for 12DCA equal to 50% of b_1 . This large SE is explained by the underlying data coming from one time series only. This does not necessarily invalidate the T_1 estimate for 12DCA, however, it simply means that the T_1 estimate for 12DCA has a wide confidence range, but is still the best estimate.

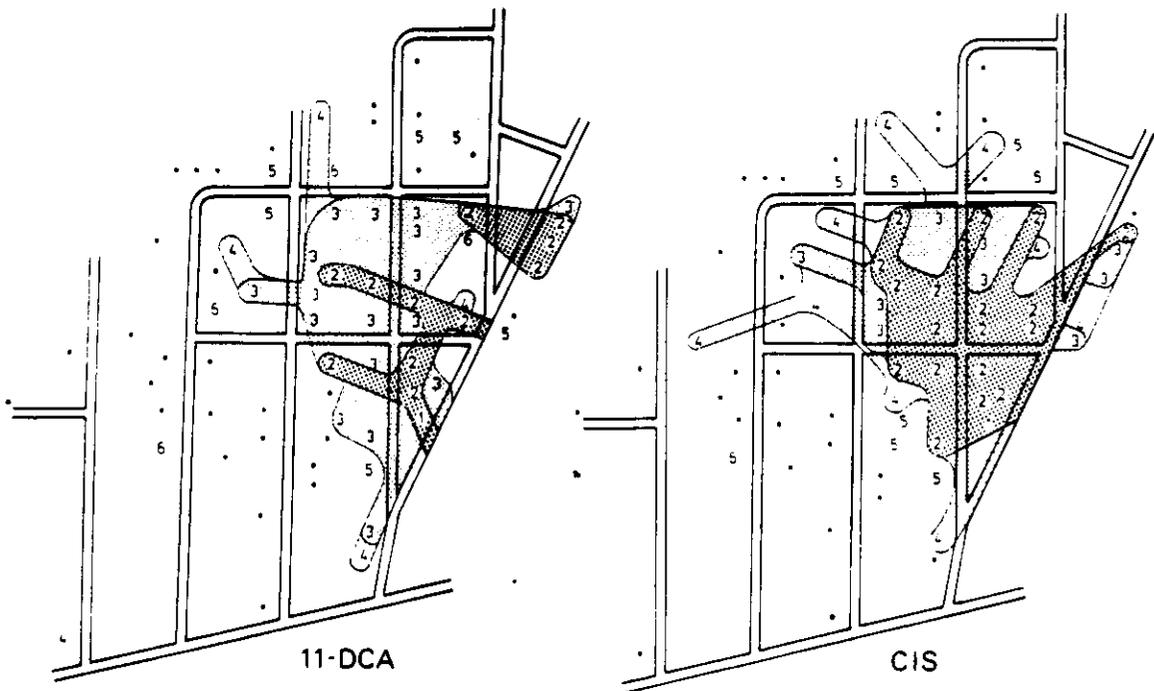


Figure C.3 Distribution of private wells west of the source of the VOC spill, grouped according to the T_1 estimate for 1,1-dichloroethane (11DCA) and cis-dichloroethylene (CIS). The T_1 estimates are grouped in periods of 1 year, the periods are coded 1 to 6. The shaded areas comprise wells with T_1 in the same period, starting Period 1 = April 1977 to March 1978, and continuing with Period 2 (April 1978 - March 1979), etc. to end with Period 6 (April 1982 - March 1983). *Uncoded* wells had no VOCs at the *time of sampling* (mostly sampled in 1982). The contour lines of the shaded areas are for illustrative purposes only, although they may approximate the actual path of progression of the plume of contaminants. For clarity, the areas with codes 5 and 6 have not been shaded.

The function $C = f(T)$ can be used for estimating the total accumulated exposure (TAE) in the period T_1 to T_s , indicated by the shaded area under the curve (Figure C.1). With individual information on wells and the period of residence in a dwelling with a contaminated well, TAE provides an individual exposure value that can be used in the analysis of epidemiologic studies. For $C_{low} = 1$ ppb and $T_{low} = T_1$, Equations #1 and #2 are written as follows:

$$\begin{aligned} C_s - 1 &= b_1(T_s - T_1) && \text{for } C = T \text{ models} \\ \log(C_s) - \log(1) &= b_1(T_s - T_1) && \text{for } \log C = T \text{ models} \end{aligned}$$

Integrating the area under these curves from T_1 to T_s yields TAE equal to:

$$\begin{aligned} \text{TAE} &= (C_s + 1)(T_s - T_1)/2 && \text{ppb-months} && \text{for } C = T \text{ models} && \text{(Eq 5)} \\ \text{TAE} &= (\exp(b_1(T_s - T_1)) - 1)/b && \text{ppb-months} && \text{for } \log C = T \text{ models} && \text{(Eq 6)} \end{aligned}$$

Using the confidence range of b_1 for $\log C = T$ models described above, the lower limit of TAE was 11-27% lower than the point estimate, with an upper limit (for VOCs other than 12DCA) of 14-61% higher than the point estimate. The widest confidence range was observed for DCE and TCE. The models for these were derived from three and two city wells respectively. This limited the number of explanatory variables (other than T), allowed in the multivariable regression models, to two for DCE and one for TCE. The upper limit for 12DCA was high, consistent with the large standard error of b_1 for this compound.

Table C.4 Range of predicted T_1 for private wells in months since January 1, 1970 and in calendar time.

VOC	Number of Wells Affected	$C = T$ based T_1		$\log C = T$ based T_1	
		Earliest	Latest	Earliest	Latest
11DCA	52	- 95 Aug 62	168 Dec 83	94 Oct 77	163 Jul 83
12DCA	48	- 333 Apr 42	152 Aug 82	87 Mar 77	152 Aug 82
DCE	34	- 234 Jul 50	149 May 82	88 Apr 77	149 May 82
CIS	66	- 1322 Nov 1849	166 Oct 83	92 Aug 77	155 Nov 82
PCE	37	43 Jul 73	163 Jul 83	108 Dec 78	163 Jul 83
TCA	44	102 Jun 78	173 May 84	113 May 79	173 May 84
TCE	43	- 578 Nov 21	158 Feb 82	100 Apr 78	152 Aug 82

C.4 Discussion

A large number of factors govern groundwater movement. Most cannot be measured directly, but some may readily be inferred from associated measurable features such as porosity, hydraulic conductivity and gradient, and the magnitude of the recharge from and discharge to the aquifer. These features are used to describe the flow in aquifers and to understand more basic characteristics of the soil or rock matrix and the water that flows through the aquifer. For instance, hydraulic conductivity reflects the friction and surface tension between soil or rock particles and water molecules, which, in turn, depends on the size, shape, and distribution of particles, as well as properties of the water (7).

An accurate forward prediction of groundwater movement requires a large number of measurements in the study area. Flow models are often applied to manipulate the large number

of variables. These models predict the movement of water, not of chemicals, but nevertheless have been used to predict movement of chemicals (8). Other, more complex models are available to simulate the transport of chemicals in a flow system. Some of these have been used successfully (9) or are still in an experimental stage (10, 11), and might theoretically be applied to the retrospective estimation of $C = f(T)$ and T_1 , but require extensive data on hydrogeologic parameters. In a typical setting where groundwater contamination was found accidentally, such as in Battle Creek, these data simply do not exist. Assuming that current parameters also prevailed in the past decade would certainly lead to erroneous results, as some parameters did change over time, e.g., the chemical gradient and the flow direction.

Because of their general concept, these prospective models are applicable to a wide variety of scenarios, but cannot discriminate between individual wells which are as close to each other as in a well field or in a residential area. Accordingly, estimates of T_1 and $C = f(T)$ inevitably refer to relatively large areas or to groups of wells rather than individual wells. This is not because the concept of the models is necessarily deficient, but because today's technology for estimating hydrogeological parameters is essentially a macroscale technology. Dealing with an individual well amidst an entire field of wells in the same aquifer would require microscale parameters. Estimating $C = f(T)$ and T_1 for individual wells would require not only hydrogeologic tests for each single well, but also the assessment of chemical-specific aquifer and soil properties still unknown. For instance, Goltz and Roberts (10) suggested the presence of immobile water compartments in an aquifer. If these exist, however, there is no method yet to assess their location and size. Neither are the causes of the immobility known, nor whether and when such compartments may become mobile again.

Questions may arise about how aquifer properties relate to T_1 and $C = f(T)$, as expressed in the Equations 1 through 4, but the retrospective $C = f(T)$ method described in this paper does not require that these questions be answered. The method is not based on hydrogeologic principles, and should not be interpreted as a mathematical hydrological flow model. It is a pragmatic and empirical approach, using statistical tools to evaluate *observed* changes in C over T . Its estimates can often be validated by observed data. The method recognizes that factors governing the movement of a chemical along the aquifer will ultimately be reflected in $C = f(T)$. The slope reflects the speed with which the chemical moves through the aquifer after being slowed down, accelerated, or detoured by the various characteristics of the aquifer, the well, and the chemical.

When a concentration gradient is created by a chemical spill, the chemical inevitably travels to locations lower on the gradient, resulting in increasing C_s in wells. It is inevitable that the movement is determined by time and characteristics of the wells and the aquifer, some of which may be known. A combination of flow, diffusion, and sorption factors, as described by Cameron and Klute (11) may explain the monotonic nature of $C = f(T)$, but not why neighboring wells have different slopes for the same chemical (see Table C.2). Acceptance of the existence of immobile compartments in the aquifer, as proposed by Goltz and Roberts (10), would provide a reasonable explanation for these different slopes, however. Immobile compartments would force the water flow to circumvent the area, resulting in increased values for dT_1 and dT_2 (Figure C.1), and thus in a later T_1 and lower slope. They also explain why some wells were contaminated and others hardly or not at all, although they were in the pathway of the plume of contamination.

The adsorption/desorption ratio eventually influences dT_1 and shapes the CT -curve. VOCs were stored and distributed at different times and volumes, depending on supply and demand, and not all tanks were corroded and leaked at the same time and at the same rate. This explains why different chemicals showed up at different times in the samples of the same well at different times, and, as a result, reached different peak concentrations.

The strategy of developing $C = f(T)$ from city wells showing increasing C over T circumvents most of the issues discussed above. Some essential underlying assumptions are necessary, however:

- 1) The private wells experienced increasing C over T at the time of sampling;
- 2) City wells selected for an increasing C over T will better represent contaminated residential wells than will a hypothetical average of clean and polluted city wells;
- 3) Private wells, once contaminated, will behave the same as city wells, that is, show the same $C = f(T)$. There are some arguments in support of these assumptions.

First, Table C.3 shows that the two private wells, tested more than once between 1981 and 1984, did show increased levels in the second sample, with CIS as the only exception. However, the decline of CIS most likely reflects a fluctuation rather than a declining curve, since large fluctuations in the CIS levels of city wells were commonly observed. Second, the assumption that private wells, when contaminated, behave as do the city wells is supported by the close agreement between estimated (modeled) and observed concentrations in these private wells. Third, as the withdrawal by private wells is minimal relative to that by city wells, the groundwater flow in the residential area is determined by the activities in the city well field (4, 5). Hence, if city wells showed increasing VOC-levels under prevailing flow conditions, it is reasonable to infer that private wells in the same period with the same flow conditions would also show increasing VOC levels. Fourth, to some degree, an increasing trend in VOC concentration was also apparent in most of the other contaminated city wells not selected for modeling. Finally, nearly all private wells were tested in 1981 and 1982 (three were tested in 1983), and none of the city wells had shown a clear decline in VOC levels before 1984.

Not all wells displayed the expected increase in C over time. This may be explained by factors other than distance, immobile aquifer compartments, or differences in sorption properties. Pumping causes a "cone of depression" in the aquifer through a pressure gradient, resulting in water from all sides being drawn to the pump. At the side of the chemical spill, the concentration will be higher than on the other sides of the cone. The 11 city wells not polluted, or barely so, were either in the vicinity of the river, or in the northern rim of the well field (Figure C.2). In other words, large supplies of clean water dominated in the cones of depression of these wells. Wells may also be affected by pumping of neighboring wells if they are within the cone of depression of those wells.

Still other assumptions are common to all regression techniques. Since our data base consisted of time series, autocorrelation may pose a source of errors (12). However, plots of residuals versus time showed a random scatter around zero over the time gradient, evidence that the models were not complicated by autocorrelation. Homogeneity of slopes of individual wells was assumed when the time series were pooled to derive a multivariable model. Testing for heterogeneity of slopes (13) showed that only one well (# 38, Table C.2) appeared to have a significantly different (steeper) slope, but only for PCE and 11DCA. However, deleting that well from the pooled data would have resulted in earlier T_1 estimates by only 0 to 2 months. This is too small a change to justify exclusion of the well from the pooled data.

$\log C = T$ models were found to have a better overall fit to the observed data than the models based on $C = T$. $\log C = T$ models proved better in another aspect as well: errors in estimating T_{low} or T_1 did not much depend on the contamination level, an important feature in view of the much wider range of C_s in private wells than in city wells (Table C.1). Using $\log C = T$ models, the range of observed deviations of estimated from observed T_{low} in city wells of 1 month to 2 years is considerably narrower than the range of errors in assessing the date a disease was diagnosed. As reported in the body of the Report, the study showed differences in

the reported date of disease onset of up to one decade, even for well-known diseases such as diabetes.

Evidence of the reasonable degree of accuracy in the estimation of T_1 may also be construed from the velocity of the groundwater. A velocity of 1-4 feet/day or about 1000 feet per year was measured (4). Well #38 in the northern rim of the well field (Figure C.2) had an observed T_1 varying from June 1982 to May 1983 for TCA, 1,1-DCA, and DCE. Given a straight distance of 1 mile from the source of the spilled chemicals, it can be concluded that water from this source would require 5 years to reach the well and, thus, that VOCs, if traveling at the same speed as groundwater, could have entered the aquifer at the spill site in approximately mid 1977 to mid 1978 to reach well #38 by the above dates. Based on similar data on observed T_1 values for various VOCs in wells #32 and #33 in the southern rim of the field, VOCs could have entered the aquifer between mid 1977 and the end of 1979. Thus, if VOCs had started their journey through the aquifer between mid 1977 and the end of 1979, they would have reached the closest private wells about 1200 feet northwest of the source sometime in 1978. This closely agrees with the T_1 estimates from $\log C = T$ models for the private wells of 1977 - 1978 (see Figure C.3). This way of estimating when leaking could have started the earliest is admittedly crude, but as it is not related to the $C = f(T)$ approach, the results certainly add to the credibility of the outcomes of $\log C = T$ models. In contrast, $C = T$ models resulted in T_1 values many decades earlier than possible, especially if the sample concentration was over 50 ppb.

The potential impact of a great number of modifications of $C = f(T)$ method were investigated, many of which were of a statistical nature. Examples of such variations include expanding the number of wells (or even pooling all wells without selection) to provide the data for slope estimation, regressing T on C rather than C on T , replacing the XY coordinates with the direct distance and the angle of the well relative to the point source, a counter-clockwise rotation of the X and Y axes by 30 degrees, deletion of influential outliers in the time series, pooling sampling results weighted by the inverse of the standard error of the slope of the wells, and regression models with interaction between T and the coordinates (to account for possible secular trends). None of these variations improved the results with regard to the difference between estimated and observed T_{low} , the ultimate parameter of accuracy, although in some instances a significant increase in the model r^2 was obtained. On the other hand, the fit of the estimated T_1 or T_{low} to the observed values was worse.

The conclusions that can be drawn from these failures are that pooling wells without applying proper selection criteria is doomed to result in a heterogeneous population of wells, rendering any modeling meaningless. Well characteristics alone are incomplete predictors of chemical movement in an aquifer. One modification, using water withdrawal parameters as additional regressor variables in the statistical model, did improve the results, however. In this version of $C = f(T)$, the number of pumping hours (H) per month, the pump capacity G of a well, and the product of these two variables (HG), which is the monthly amount of water withdrawn from the aquifer, were used as additional regressors. In addition to an improved r^2 , the regression slopes increased slightly, resulting in T_1 estimates later by two months or less. This was expected on the basis of theoretical considerations fitting the hypothesis on factors governing T_1 and b_1 . Since H is a true time-dependent variable and not a well constant, actual data on H are required to use the improved model. Unfortunately, as no information on H was available for private wells, this model could not be applied to these wells. A possible explanation for the limited success of this version of $C = f(T)$ is that the compound effect of the interaction of cones of depression of neighboring wells, and the relation of these cones to the proximity of the river and the main axis of the groundwater flow, are complex matters requiring procedures more sophisticated than simply adding the parameters to the model.

Clearly, the method described cannot, and should not, be used to prospectively predict whether, when, and to what degree water from a given well may become contaminated with a

chemical of interest. On the other hand, the results do show that, if no significant alterations in the local management of an aquifer have been made in the recent past, the results of a current, simple, monitoring program for nearby wells can be applied to unmonitored wells to yield data on the estimated time that the chemicals reached the unmonitored wells, and how the chemical concentrations changed over time. Evidence has been provided of reasonable accuracy of the results, which could not have been obtained by any of the existing hydrologic flow models. The $C = f(T)$ approach is not presented as the ultimate solution for retrospectively estimating the movement of chemicals in groundwater. It is rather a pragmatic tool for quantifying exposure for epidemiologic purposes superior to the conventional, but incorrect, approach of assuming a constant exposure level throughout the period of residence in the area. There is a good possibility that the described method is applicable to other sites with a similar scenario of the presence of some monitoring data for related wells and a point source.

C.5 The Source of Pollution

This study was not intended to prove that the alleged source of VOCs, as shown in Figure C.2, was indeed the only source contaminating the aquifer. Given the assumption, that there was no other source, the results of the modeling were surprisingly accurate. It was known from the outset that two other sources of pollution might be present, although probably of lesser magnitude. Therefore, one of the modifications of the model tested was to use a hypothetical source 4 grids west of the primary source indicated in Figure C.1 (a solvent handling site belonging to the same company). This impaired the model fit to such an extent that this second source can readily be dismissed as a source with significant impact on the concentration levels. No modeling was attempted to fit the model to the third location, east of the city well field at the railroad yard. As discussed in this study, the model depends on the ability to locate the source rather precisely. The yard is too large, relative to the distance from individual city and private wells, to allow the use of an estimate of the central point of spilling small enough to fit in one grid. An indirect argument against an important role of the yard is the finding that samples from the city wells between the railroad and wells 27, 28, and 29 did not show significant levels of VOC contamination.

The discussion above merely emphasizes the absence of evidence that secondary sources played a substantial role. For instance, a comparison of the kind, frequency, and levels of the VOCs found in the city wells with those in the private wells west of the source does reveal a feature insufficiently explained by our model. CIS and 12DCA were present in private wells at levels one order of magnitude higher than in city wells, while the frequency of occurrence (the percentage of wells polluted with these two chemicals) was also much higher. The private wells are closer to the source, following the hypothesized pathway of the plume of contaminants, and a higher concentration is expected. But this large difference has not been found for the other chemicals. It is, thus, not ruled out that there was a secondary source of CIS and 12DCA closer to the residential wells than to the city wells, and probably outside the main flow from the main source to the Battle Creek River.

There is a scenario in which this difference in the contamination pattern can be explained while maintaining the single source hypothesis. In this scenario, the spilling of CIS and 12DCA ceased at or around 1981, when the contamination of the aquifer was detected. Since that time, the city wells were protected by increased pumping-to-waste of "barrier" wells. This would accelerate the progress of the plume of VOCs to the residential wells, and peak concentrations in those wells could occur, but not in the city wells. This hypothesis is speculative, although it would fit the data, and does not require the presence of a second or third source of pollution, for which there is no support in the data. Support for, or evidence against, the above hypothesis might have been obtained if logbook information (data and volumes) were made available on the transportation and handling of individual VOCs.

C.6 Application of the Results to the Battle Creek Health Study

The point estimate of the slope of the $\log C = T$ model (b_1), derived from stepwise regression models using Equations 4 and 6, has been used to obtain individual estimates of T_1 and the accumulated exposure TAE for use in the epidemiologic analysis of the potential health effect of VOC exposure. It needs to be stressed that these estimates are site- and scenario specific, and cannot and should not be used for any other site or chemical. The estimates of b_1 used for calculating T_1 and TAE are listed below. For individual estimates of TAE (area under the CT-curve), the dates of moving to and out of the house with the contaminated well, and the date that people stopped drinking VOC-contaminated water, must be known.

Table C.5 Estimations of b_1 for the seven VOCs.

Compound	Abbreviation	b_1
1,1-dichloroethane	11DCA	0.1034
1,2-dichloroethane	12DCA	0.1048
1,1,1-trichloroethane	TCA	0.1536
1,1-dichloroethylene	DCE	0.0795
cis-1,2-dichloroethylene	CIS	0.1667
1,1,1-trichloroethylene	TCE	0.1434
1,1,2,2-tetra or perchloroethylene	PCE	0.1493

C.7 Conclusions

- 1) Under the conditions of the study area, well monitoring data can be used to estimate the changes in the concentration of a chemical in well water as a function of time, $C = f(T)$.
- 2) A chemical-specific $C = f(T)$, developed from monitoring data, can be applied to nearby contaminated wells lacking such data, for retrospectively estimating the time T_1 when the contamination started, and how the VOC-levels changed since T_1 , provided that no major changes in the local aquifer management took place in the recent past.
- 3) The results of fitting $C = f(T)$ to the data of the few wells that provided an opportunity for validation, suggest that the assumed source of pollution is indeed the major source of most or all of the VOCs studied.
- 4) Under the conditions of the study site, the best of the models studied is:

$$\log C = a + b_1 T_1 + b_2 X_i$$

In this equation, X_i represents well characteristics, such as distance from the source (coordinates, or straight distance and angle), depth, and pump capacity, hours of pumping, and the amount of water withdrawn. The specificity to the chemical of interest can be achieved through a stepwise version of the regression analysis.

- 5) Information on when individuals moved into or out of the dwelling with the contaminated well, and when they stopped using the water can be used in the equations to yield individual values of total accumulated exposure.

C.8 Summary

For a proper analysis of the potentially causal relationship between exposure to volatile organic chemicals (VOCs) in drinking water and health events, it is essential to know T_1 , the time when exposure started, and $C = f(T)$, which is the change of the VOC concentration C as a function of time T , and the total accumulated exposure (TAE) to VOCs to which an individual was exposed. In the typical situation of incidentally detected pollution of groundwater, no such information is available. This paper describes the development of a method for estimating T_1 , TAE and $C = f(T)$, as part of an epidemiologic study of the health effects of VOC contamination of aquifers serving public and private wells. Pooled test results of city wells, monitored since 1981, provided the data base for developing a statistical model for estimating $C = f(T)$. This model was then applied to private wells, for which the data of only one water sample were available, to retrospectively estimate their T_1 . The best fitting model was a multiple linear regression equation consisting of the natural logarithm of the VOC concentration as the response variable, with the time of sampling, the distance of the wells from the source (expressed as coordinates), the well depth, and the well capacity and actual output as determinants. The TAE was calculated by integrating the area under the time-concentration curve.

C.9 Acknowledgements

We are indebted to Messrs. Arthur W. Bloomer, Adrian J. Oudbier, and Jon W. Bloemker of the Michigan Dept. of Public Health, and Mr. Theodore R. Havens from the Calhoun County Health Department for their assistance in compiling the necessary data. We thank Mr. Norman G. Grannemann, hydrologist, U.S. Geological Survey, for his invaluable advice with regard to the hydrologic aspects of this paper.

C.10 References

1. Littlefield, N. A., Farmer, J. H., Gaylor, D. W., and Sheldon, W. G. Effects of dose and time in a long-term, low dose carcinogenic study. In: *Innovations in cancer risk assessment*. (J. E. Staffa and M. A. Mehlman, eds.), Pathotox Publishers, Park Forest South, 1979, pp. 17-34.
2. Brown, C. C. and Chu, K. C. Implication of the multistage theory of carcinogenesis applied to occupational arsenic exposure. *J. Natl. Cancer Inst.* 70: 455-463(1983).
3. CH2M - Hill. *Remedial action master plan, Verona Well Field, Battle Creek, Michigan*. Report to the U.S. Environmental Protection Agency, 1983.
4. Grannemann, N. G. and Twenter, F. R. *Geohydrology and Groundwater Flow at Verona Well Field, Battle Creek, Michigan*. U.S. Geological Survey report 85-4056(1985).
5. CH2M - Hill. *Technical Memorandum Phase I: Drilling and Soil Sampling Verona Well Field, Battle Creek, Michigan*. REM/FIT report C11185 to the U.S. EPA, 1985.
6. SAS Institute. Stepwise procedure. In: *User's Guide: Statistics*, 5th ed. SAS Institute, Cary, 1985, pp. 763-774.
7. Shervani, J. W. Groundwater development. In: *Elements of Water Supply and Wastewater Disposal*. (G. M. Fair, J. C. Geyer, and D. A. Okun, eds.), 2nd ed., Wiley & Sons, New York, 1971, pp. 112-162.
8. Bond, F. W., Cole, C. R., and Sanning, D. Evaluation of remedial action alternatives: demonstration/application of groundwater modeling technology. In: *Proc. Natl. Conf. Management of Uncontrolled Hazardous Waste Sites*, HCMRI, Silver Spring, November 1982, pp. 118-123.
9. Konikow, L. F. *Modeling Chloride Movement in the Alluvial Aquifer at the Rocky Mountain Arsenal, Colorado*. U.S. Geological Survey Water Supply report nr. 2044, 1977.
10. Goltz, M. N. and Roberts, P. V. Interpreting organic solute transport data from a field experiment using physical nonequilibrium models. *J. Contam. Hydrol.* 1: 77-93(1986).
11. Cameron D.R. and Klute A. Convective-dispersive solute transport with combined equilibrium and kinetic adsorption model. *Water Resources Res.* 13: 183-188(1977).
12. Neter J., Wasserman W., and Kutner M.H. *Applied Linear Statistical Models*. 2nd ed., 1985, pp. 444-465. Richard D. Irwin, Homewood Ill.
13. Freund R.J. and Littell R.C. SAS for linear models. In: *Heterogeneity of Slopes*, SAS Institute, Cary, 1985, pp. 200-205.

Appendix D

**Mortality and Morbidity Rates in Michigan
for the State and Selected Counties and Minor Civil
Divisions Involved in the Battle Creek Health Study**

**Maurice C. Barone
Center for Health Statistics
Michigan Department of Public Health
Lansing, Michigan**

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

**Janet T. Eyster, PhD
Center for Health Statistics
Michigan Department of Public Health
Lansing, Michigan**

D.1 Introduction

The Battle Creek Health Study entailed the estimation of levels of exposure to VOCs and the frequency of health events of small populations near the Verona Well Field and Springfield, both adjacent to Battle Creek City, Calhoun County, Michigan, and a neighborhood in Dowagiac City, Cass County, Michigan. A statistical analysis was made of associations between exposure and disease frequency. In order to provide a proper reference, the study included populations in Calhoun County (Ceresco, Emmett Township, and Pennfield Township), Battle Creek City, and Cass County. It was necessary to estimate mortality and morbidity rates in these areas to provide background data for the study areas. Increased rates may be caused by conditions prevailing at the neighborhood or at more extensive levels. In case of the former, one would not expect that excess disease will show up in the rates of State or County populations. Therefore, available State and County-wide data would need to be broken down to the level of the smallest possible population unit.

The health study involved an evaluation of possible adverse health effects of contamination to volatile organic chemicals (VOCs) in drinking water. Given the toxicity profile of the VOCs, the most likely to be affected are the rates of diseases of the liver, kidney, nervous system, reproductive events, and cancer. Although these diseases are known to occur in animals and humans exposed to high VOC levels, the effect of relatively low doses involved in water contamination is unknown.

State-wide mortality data have frequently been used as a standard reference when evaluating reports on mortality clusters or excess mortality in subgroups of the population. This requires some assumptions to be made:

- 1) The study population is representative of the reference population.
- 2) The demographic information stored in the reference data base is accurate for age, sex, race, and residency.
- 3) The reference data base is accurate as to the recording the primary cause of death, and allows access to all other diseases on the death certificate.
- 4) The reference data base provides an accurate record of the actual diseases present at the time of death, and of the actual causes leading to the fatal outcome.

Obviously, the first assumption holds for any study in which a sample of the target population is investigated. However, populations living near a chemical dump site or with a certain occupational exposure are rarely, if ever, representative of the state or county population. Often, some impression may exist of differences between sample and reference. But in virtually all situations, the magnitude and extent of these differences are unknown, rendering the value of inferences from studies comparing specific populations with state or county populations dubious. The second assumption is important because populations are compared on the basis of demographic parameters. Errors can be made when transferring data from the death certificate, or if the data on the certificate are inaccurate by themselves. The fourth assumption touches upon the problem of interpreting the logic in the sequence of diseases reported on the death certificate.

Whether diabetes is the cause of death rather than kidney failure or bacterial septicemia following infection of skin ulcers, requires a medical expertise usually not available in the typical offices of vital statistics. This can be remedied, however, by methods of coding that allow access to any disease listed on the certificate. The last assumption highlights a major problem that has been studied extensively, but is still unsolved. Persons filling in the death certificate are not always familiar with the medical history of the deceased, the clinical status preceding the

death, or the autopsy findings which are only available some time later. The routine character of the death certificate is almost a guarantee that only the diseases considered to be related to the cause of death, or considered to be important, will be recorded,

Mortality data are a poor reflection of the morbidity pattern in populations. Most diseases are not fatal or may not have contributed to the cause of death. The proper indicator of morbidity would be the incidence rate. Except for cancer, population-based incidence rates for noninfectious diseases are rare. With regard to cancer, the only population-based cancer registry in Michigan, that of the Detroit Metropolitan Area, does not cover the Battle Creek Health Study population. Assumptions analogous to that for mortality data also apply to hospital discharge data. An assumption specific to hospital data is that the disease of interest requires hospitalization, and that the admission rate is equal for the study population and the reference population. It is obvious, however, that hospitalization depends not only on the severity of the disease of concern, but also on the socioeconomic status and on local differences in medical practice. Further, people are usually hospitalized more than once for the same disease. Thus, discharge data will certainly show patterns different from incidence or mortality data. It is, however, reasonable to expect that general trends in morbidity and mortality will also be reflected to some extent in discharge data.

The above discussion is directly relevant to the current epidemiologic study. The study populations mentioned earlier are predictably not representative for the total population of either Calhoun or Cass Counties. Unfortunately, the data available, that of statewide hospital discharge files and death certificates, cannot be brought down to the size of the exposed and reference populations. Calculating mortality data from civil divisions smaller than a county requires ascertainment of the accuracy of demographic characteristics of the decedents. The need for verifying data from interviewing the next of kin with the Vital Statistics data base demands that the latter can be relied upon as a decisive source of information. Investigation of the accuracy of the Vital Statistics data base of the State of Michigan was made part of Phase II of the Battle Creek Health Study. To serve as a reference for the outcomes of that study, mortality and hospital discharge rates were calculated for the areas involved in the epidemiological study. The following chapters discuss the results of this investigation.

D.2 Mortality Rates

In the State of Michigan, death certificates are prepared by physicians and funeral directors, and filed with 114 local registrar offices. After reviewing for accuracy and completeness, local registrars forward the certificates to the Office of the State Registrar and Center for Health Statistics at the Michigan Department of Public Health (MDPH). Mortality data files are developed through manual coding and key entry of demographic and cause of death entries on death certificates. Coding and key entries are subject to qualification and sample verification procedures. Death data are edited in conformance with standards established by the National Center for Health Statistics (NCHS). Multiple causes of death are evaluated to produce a standardized determination of the underlying cause of death, using the NCHS Automated Classification of Medical Entities (ACME) programs. Edited data are sent to NCHS where similar editing occurs in conjunction with data quality sampling. These data files are then a source for standard administrative and statistical reports on a monthly, quarterly and annual cycle as well as for "as needed" summaries and reports.

A study of mortality rates for the Battle Creek area, following the detection of the contamination of the Verona Well Field, showed an excess mortality for a number of diseases, as shown in Table D.1. An analysis of the quality of the data revealed inaccuracies in the coding of the information relating to the minor civil division recorded on or transferred from the death certificate. When the civil division of the decedent's address was verified using maps or other

materials, up to 25 percent of the addresses appeared to be located outside the recorded city, township or village. The inaccuracies included errors in recording the minor civil division on the death certificate, and coding errors as well. Attempts were made as a part of Phase II of the Battle Creek Health Study, to correct the errors death data in the Dowagiac and Battle Creek areas, for the period 1970 through June, 1984. Street addresses, recorded on death certificates from Cass and Calhoun county residents, were extracted for determination of the minor civil division and for recoding. In the few cases with insufficient street address information, the original recorded minor civil division was maintained.

State and county population estimates, obtained from the Michigan Department of Management and Budget, are consistent with Michigan population estimates from the U. S. Bureau of the Census. Population estimates by age, race and sex are developed by the Office of the State Registrar and Center for Health Statistics using proportional extrapolation from the 1970 and 1980 censuses. The subgroup proportions were then applied to the state and county totals to obtain population estimates. It should be noted that for the 1980 census, the modified race distribution prepared by the U.S. Bureau of the Census was used. This modified race distribution is consistent with race categories used in the 1970 census, but differs from the published 1980 population figures.

Minor civil division population estimates were prepared by the Office of the State Registrar and Center for Health Statistics, using population figures for the 1970 and 1980 censuses published by the U.S. Bureau of the Census. Age, race, and sex estimates for Battle Creek City, Springfield and Dowagiac City were derived as described above. For the minor civil divisions Pennfield and Emmett Townships, the 1970 age distribution was estimated from the 1980 census, as no age distribution information was available for the 1970 populations. For the 1980 census, it should be noted that an unmodified race distribution was used for the minor civil division population estimates.

The underlying cause of death is that condition considered to have given rise to the chain of events leading to death. An individual death was classified as having only one underlying cause. Beside the underlying cause of death, death certificates may have other conditions listed as secondary or related causes of death. In the current study, mortality rates for the period 1970 through June of 1984 were based on deaths with any mention of the condition of interest as either an underlying or related cause of death. Therefore, a single death may be included in the mortality rates for more than one disease.

For this study, mortality rates have been calculated for whites only, since the percentage of whites in the Phase II study population is 94%. Mortality rates for the counties and minor civil divisions were tested against the State rate to determine if they were significantly higher using a procedure described by Bailar and Ederer (1). Except for Table D.1, mortality rates have been directly adjusted for age, using the Michigan 1980 population as the standard, to allow comparisons be made between subpopulations. Table D.1 shows the results of the pilot study as indirectly standardized mortality ratios (SMRs).

SMR refers to the ratio of deaths observed in the study area over deaths expected from U.S. age-specific rates for 1978 and estimated age-specific populations for Battle Creek City and Calhoun County, 1978. The data have *not* been corrected for errors.

Table D.1 Standardized mortality ratios (SMRs) for Battle Creek City and Calhoun County for ten leading causes of death during 1976-1980.

Cause of Death	Battle Creek City	Calhoun County
Heart diseases	1.306 **	0.986
Malignant neoplasms	1.215 **	0.973
Cerebrovascular disease	1.198 **	0.996
Accidents and adverse effects	0.972	0.925
Chronic obstructive and allied pulmonary diseases	1.614 **	1.141 *
Pneumonia and influenza	0.829	0.731
Diabetes mellitus	2.135 **	1.664 **
Atherosclerosis	1.261	1.198
Chronic liver disease and liver cirrhosis	2.314 **	1.119
Suicide	1.441 *	0.777
All other	1.240 **	0.902

- * = significantly greater than the Michigan rate at $p < 0.025$ (1-tail test)
 ** = significantly greater than the Michigan rate at $p < 0.005$ (1-tail test)

D.2.1 Results

Preliminary analysis of uncorrected death data for underlying cause of death from Battle Creek City for the years 1976 through 1980 indicated that the city population had significantly greater numbers of deaths than expected based on 1978 U.S. death rates for 7 of the 10 leading causes of death: diseases of the heart, malignant neoplasms, cerebrovascular diseases, chronic obstructive lung diseases and allied conditions, diabetes mellitus, chronic liver disease and cirrhosis, suicide, and "all other causes" (Table D.1). The Verona area and Springfield are adjacent to, but not part of, Battle Creek City. The Verona area encompasses adjacent parts of Emmett and Pennfield Townships. Springfield is a separate city. Death data for underlying cause of death from Calhoun County for the same period showed that Calhoun County had significantly greater numbers of deaths than expected for two of the ten leading causes of death: chronic obstructive pulmonary diseases and allied conditions and diabetes mellitus.

Tables D.2 (males) and D.3 (females) show directly age adjusted mortality rates for deaths with any mention of 12 selected diseases for residents of Michigan, Cass and Calhoun Counties and five minor civil divisions: Dowagiac City (Cass County), Battle Creek City, Springfield, Emmett Township and Pennfield Township (all Calhoun County) for the time periods 1970-74,

1975-79, and 1980-84. The rates were calculated from death certificate information after correction of inaccuracies in the recording of the minor civil division (cities, townships and villages). While a scattering of statistically significant rate increases is apparent, notable patterns emerge in only a few instances.

The Battle Creek City all-cancer mortality rate for deaths with any mention of a malignant neoplasm was higher than the State rate. This increase was statistically significant except for females in 1970-74. Liver disease rates were higher for Battle Creek and Dowagiac males in all periods, but the difference was statistically insignificant for Dowagiac in the periods before 1980. In contrast, the State rates for liver diseases dropped consistently over time. Alcohol related deaths were more frequent in Battle Creek for both sexes, but statistical significance was not reached for females in the periods since 1975. There was an unexplained steep peaking of the rate for males in 1975-79.

In comparing Tables D.2 and D.3 with Table D.1, it should be noted that the U.S. population was the reference in calculating the SMRs in Table D.1. The Michigan population was used in Tables D.2 and D.3. Although this deviation may have some impact on the differences between subpopulations, these differences are quite small compared to differences caused by using a data base which was uncorrected for errors and inaccuracies in the residency data for Table D.1. The increase in diabetes mortality (Battle Creek 100+ %, Calhoun county 60+ %), compared to the State (Table D.1), disappeared in Tables D.2 and D.3. In fact, rates became lower, although statistically insignificant.

The data in Tables D.2 and D.3 do not support the hypothesis that increased liver disease mortality may have resulted from drinking VOC contaminated water. First, the increased rate for liver diseases found in the pilot study (Table D.1) was not confirmed for females. Second, in the areas affected by contamination (Emmett and Pennfield townships), mortality with any mention of liver disease was even lower than in Calhoun county and Battle Creek, except for Pennfield females in the period 1975-79. Third, the increased mortality rates for liver disease can sufficiently be explained by the concurrently higher rate for alcohol related diseases.

Table D.2 Number and rates for deaths with any mention of 12 selected diseases. Age-adjusted rate (reference: Michigan white 1980) per 100,000 for White Males.

Disease	Area	1970 - 1974		1975 - 1979		1980 - 1984	
		N	rate	N	rate	N	rate
All malignant neoplasms	Cass County	219	220.4	229	210.4	253	220.9
	Dowagiac	34	253.0	29	215.7	41	288.1
	Calhoun County	644	206.8	692	212.8	773	225.7
	Battle Creek	226	253.7 *	215	262.3 *	212	277.6 *
	Springfield	19	207.3	31	283.8	27	208.1
	Emmett Townsh.	44	170.6	51	191.2	70	252.8
	Pennfield Townsh.	36	170.8	35	160.3	43	198.2
	Michigan State	39390	219.2	42154	225.2	44965	231.7
All cancer of esophagus and stomach	Cass County	11	11.2	11	9.7	11	9.5
	Dowagiac	6	46.3 *	1	8.3	1	8.6
	Calhoun County	26	8.3	37	11.3	26	7.6
	Battle Creek	6	7.1	13	16.0	9	11.7
	Springfield	0	0	2	19.0	0	0
	Emmett Townsh.	1	3.1	3	11.5	4	13.9
	Pennfield Townsh.	1	4.4	3	11.9	3	12.2
	Michigan State	2786	15.5	2631	14.1	2533	13.1
Cancer of intestines, colon, and rectum	Cass County	27	27.7	30	27.5	28	24.7
	Dowagiac	7	54.0	1	7.6	4	29.4
	Calhoun County	81	26.1	94	28.9	90	26.2
	Battle Creek	35	38.6	22	27.0	28	36.9
	Springfield	5	50.9	5	46.8	3	20.7
	Emmett Townsh.	1	3.8	7	26.7	11	39.1
	Pennfield Townsh.	4	19.1	3	15.8	5	24.0
	Michigan State	5147	28.8	5424	29.1	5442	28.0
Cancer of liver and intrahepatic bile ducts	Cass County	0	0	0	0	3	2.6
	Dowagiac	0	0	0	0	1	6.8
	Calhoun County	8	2.7	2	0.6	15	4.3 *
	Battle Creek	2	2.2	0	0	3	4.2
	Springfield	1	10.3	0	0	0	0
	Emmett Townsh.	0	0	0	0	1	3.8
	Pennfield Townsh.	0	0	0	0	1	4.2
	Michigan State	254	1.4	328	1.7	611	3.1
Cancer of bladder & kidneys	Cass County	18	18.3	16	14.5	18	15.6
	Dowagiac	2	15.0	2	14.3	0	0
	Calhoun County	52	16.5	49	15.1	51	14.9
	Battle Creek	20	22.1	16	19.8	14	16.5
	Springfield	2	24.5	2	16.6	1	6.3
	Emmett Townsh.	5	19.7	5	18.5	6	21.8
	Pennfield Townsh.	4	20.5	3	12.4	3	12.3
	Michigan State	2722	15.2	2703	14.5	2840	14.6

Table D.2 Continued: Mortality Rates for White Males

Disease	Area	1970 - 1974		1975 - 1979		1980 - 1984	
		N	rate	N	rate	N	rate
Leukemias	Cass County	18	17.9	14	13.0	9	8.0
	Dowagiac	0	0	5	36.3 *	2	13.8
	Calhoun County	27	8.5	30	9.3	36	10.8
	Battle Creek	11	12.6	8	9.4	7	9.3
	Springfield	1	11.7	0	0	3	25.1
	Emmett Townsh.	2	7.0	3	10.8	4	14.6
	Pennfield Townsh.	2	8.8	1	4.4	0	0
	Michigan State	2290	12.5	2175	11.5	2147	11.1
Diabetes	Cass County	38	38.6	44	40.5	61	53.2
	Dowagiac	6	43.5	3	21.9	13	81.6
	Calhoun County	202	65.0	188	58.0	186	54.1
	Battle Creek	71	78.9	52	62.2	59	80.6
	Springfield	6	67.3	5	44.1	10	75.8
	Emmett Townsh.	19	74.6	20	77.0	10	37.0
	Pennfield Townsh.	10	50.0	12	58.9	11	49.9
	Michigan State	13272	74.5	12004	64.4	11786	60.7
Alcohol dependency	Cass County	6	5.9	6	5.5	10	8.6
	Dowagiac	1	8.1	3	25.4	2	15.4
	Calhoun County	30	9.8	59	18.5 **	28	8.3
	Battle Creek	16	20.9 *	35	53.6 **	16	25.2 *
	Springfield	1	9.6	5	40.4 *	2	13.1
	Emmett Townsh.	4	15.6	9	30.4 *	1	3.7
	Pennfield Townsh.	0	0	2	8.9	1	4.5
	Michigan State	1960	10.6	2062	10.9	1806	9.3
Liver necrosis	Cass County	0	0	1	0.8	1	0.8
	Dowagiac	0	0	0	0	0	0
	Calhoun County	2	0.6	4	1.3	2	0.6
	Battle Creek	1	1.3	1	1.0	1	1.5
	Springfield	0	0	0	0	0	0
	Emmett Townsh.	0	0	0	0	0	0
	Pennfield Townsh.	0	0	0	0	0	0
	Michigan State	140	0.8	121	0.6	226	1.2
Liver disease and cirrhosis	Cass County	22	22.1	21	19.0	17	14.6
	Dowagiac	9	77.4	5	41.2	7	68.7 *
	Calhoun County	98	30.5	104	31.6	76	22.4
	Battle Creek	46	59.1 **	58	81.2 **	30	46.6 *
	Springfield	4	37.9	3	24.3	5	33.8
	Emmett Townsh.	7	23.6	8	27.6	3	9.4
	Pennfield Townsh.	2	8.3	2	8.1	3	14.1
	Michigan State	6977	37.8	5810	30.7	4979	25.8

Table D.2 Continued: Mortality Rates for White Males

Disease	Area	1970 - 1974		1975 - 1979		1980 - 1984	
		N	rate	N	rate	N	rate
Diseases of the pancreas	Cass County	3	3.1	1	0.9	4	3.5
	Dowagiac	0	0	1	8.3	2	21.2
	Calhoun County	19	6.2	10	3.0	11	3.3
	Battle Creek	8	10.5 *	0	0	2	2.3
	Springfield	0	0	1	7.7	2	16.7
	Emmett Townsh.	2	7.1	3	11.3	2	7.5
	Pennfield Townsh.	1	4.4	0	0	0	0
	Michigan State	818	4.5	668	3.5	579	3.0
Kidney diseases	Cass County	10	10.2	6	5.7	47	42.4
	Dowagiac	3	22.7	1	6.0	9	53.2
	Calhoun County	34	10.8	62	19.1	152	44.5
	Battle Creek	10	11.6	28	31.2 *	44	57.1
	Springfield	2	23.3	4	32.0	10	75.4
	Emmett Townsh.	3	12.5	2	7.8	7	25.3
	Pennfield Townsh.	1	4.7	3	13.0	12	58.4
	Michigan State	1935	10.8	3667	19.6	9602	49.5

* = significantly higher than the Michigan rate, $p < 0.025$ one-sided

** = significantly higher than the Michigan rate, $p < 0.005$ one-sided
(for significance testing, see Reference)

Table D.3 Number and rates for deaths with any mention of 12 selected diseases. Age-adjusted rate (reference: Michigan white 1980) per 100,000 for White Females.

Disease	Area	1970 - 1974		1975 - 1979		1980 - 1984	
		N	rate	N	rate	N	rate
All malignant neoplasms	Cass County	149	162.2	167	160.7	163	147.1
	Dowagiac	31	198.1	34	190.7	32	177.4
	Calhoun County	561	174.8	594	172.6	666	182.3
	Battle Creek	207	189.6	208	201.5 *	224	233.8 **
	Springfield	13	127.1	21	154.9	23	140.4
	Emmett Townsh.	49	209.6	45	167.9	51	173.5
	Pennfield Townsh.	22	106.3	28	130.2	38	177.8
	Michigan State	31081	173.6	33696	174.4	38078	185.4
All cancer of esophagus and stomach	Cass County	8	8.9	8	7.9	5	4.5
	Dowagiac	2	13.1	0	0	1	3.8
	Calhoun County	23	7.2	11	3.2	22	5.9
	Battle Creek	8	6.1	4	3.7	9	8.7
	Springfield	0	0	0	0	0	0
	Emmett Townsh.	3	14.2	1	4.2	2	6.9
	Pennfield Townsh.	0	0	2	8.7	2	11.1
	Michigan State	1438	8.2	1390	7.3	1326	6.4
Cancer of intestines, colon, and rectum	Cass County	23	25.3	19	18.5	27	25.0
	Dowagiac	7	41.1	4	22.1	5	19.0
	Calhoun County	92	29.0	105	30.4	89	23.5
	Battle Creek	31	26.2	37	34.3	29	24.1
	Springfield	1	11.0	2	17.2	1	4.1
	Emmett Townsh.	9	43.0	6	23.1	13	42.9
	Pennfield Townsh.	2	9.9	8	37.7	3	13.9
	Michigan State	4881	28.0	5021	26.2	5342	25.9
Cancer of liver and intrahepatic bile ducts	Cass County	0	0	0	1.0	0	0
	Dowagiac	0	0	0	0	0	0
	Calhoun County	2	0.6	2	0.6	7	2.0
	Battle Creek	0	0	0	0	3	3.6
	Springfield	0	0	0	0	0	0
	Emmett Townsh.	0	0	0	0	2	6.8
	Pennfield Townsh.	0	0	0	0	0	0
	Michigan State	128	0.7	173	0.9	462	2.2
Cancer of bladder & kidneys	Cass County	5	5.3	4	4.1	4	3.4
	Dowagiac	2	15.1	0	0	1	7.2
	Calhoun County	11	3.4	24	6.9	21	5.6
	Battle Creek	3	3.0	10	9.1	10	9.7
	Springfield	0	0	1	9.5	0	0
	Emmett Townsh.	0	0	4	15.9	0	0
	Pennfield Townsh.	1	5.1	0	0	2	11.1
	Michigan State	1158	6.6	1302	6.8	1340	6.5

Table D.3 Continued: Mortality Rates for White Females

Disease	Area	1970 - 1974		1975 - 1979		1980 - 1984	
		N	rate	N	rate	N	rate
Leukemias	Cass County	9	9.9	10	9.5	9	8.2
	Dowagiac	3	18.7	3	14.8	2	16.8
	Calhoun County	30	9.3	21	6.1	20	5.5
	Battle Creek	11	10.0	8	7.1	7	7.0
	Springfield	0	0	0	0	0	0
	Emmett Townsh.	4	14.9	0	0	2	7.2
	Pennfield Townsh.	1	5.8	0	0	2	9.6
	Michigan State	1564	8.6	1536	7.9	1587	7.7
Diabetes	Cass County	58	64.6	61	59.8	72	66.6
	Dowagiac	12	77.5	17	89.2	17	74.1
	Calhoun County	263	82.8	219	63.0	242	64.0
	Battle Creek	106	91.6	82	74.3	77	73.4
	Springfield	8	86.3	8	56.3	19	100.4
	Emmett Townsh.	21	97.5	19	74.2	27	89.6
	Pennfield Townsh.	7	36.2	7	34.5	14	71.2
	Michigan State	16532	96.2	14861	78.0	14673	70.9
Alcohol dependency	Cass County	7	7.3 *	3	2.7	2	1.7
	Dowagiac	5	36.9 *	0	0	0	0
	Calhoun County	8	2.3	11	3.3	7	2.0
	Battle Creek	3	4.0	4	5.3	6	7.9 *
	Springfield	1	7.5	1	6.9	0	0
	Emmett Townsh.	0	0	2	7.0	0	0
	Pennfield Townsh.	1	4.3	0	0	1	4.0
	Michigan State	521	2.7	547	2.8	471	2.3
Liver necrosis	Cass County	0	0	0	0	1	1.0
	Dowagiac	0	0	0	0	0	0
	Calhoun County	5	1.6	2	0.6	2	0.6
	Battle Creek	0	0	0	0	1	0.8
	Springfield	0	0	0	0	0	0
	Emmett Townsh.	0	0	0	0	0	0
	Pennfield Townsh.	0	0	0	0	0	0
	Michigan State	143	0.8	80	0.4	136	0.7
Liver disease and cirrhosis	Cass County	17	17.6	8	7.4	11	9.8
	Dowagiac	5	36.0	1	4.5	2	14.3
	Calhoun County	52	15.8	43	12.8	47	13.3
	Battle Creek	20	20.1	14	16.9	23	36.1 *
	Springfield	2	18.5	2	16.1	0	0
	Emmett Townsh.	4	15.4	7	25.8	6	20.1
	Pennfield Townsh.	2	9.4	3	14.9	0	0.0
	Michigan State	3057	16.2	2560	13.0	2370	11.6

Table D.3 Continued: Mortality Rates for White Females

Disease	Area	1970 - 1974		1975 - 1979		1980 - 1984	
		N	rate	N	rate	N	rate
Diseases of the pancreas	Cass County	2	2.3	4	3.8	1	1.0
	Dowagiac	1	6.6	0	0	0	0
	Calhoun County	11	3.4	15	4.4 *	11	3.2
	Battle Creek	8	9.4 *	6	5.4	5	6.6
	Springfield	0	0	0	0	0	0
	Emmett Townsh.	1	3.9	3	11.8	0	0
	Pennfield Townsh.	0	0	0	0	2	9.7
	Michigan State	552	3.0	497	2.6	449	2.2
Kidney diseases	Cass County	7	7.8	13	13.0	38	35.0
	Dowagiac	1	5.6	6	29.6	11	49.8
	Calhoun County	39	12.2	67	19.3	139	36.9
	Battle Creek	10	8.2	32	30.9 *	45	46.9
	Springfield	1	11.0	6	39.2	9	40.6
	Emmett Townsh.	4	18.1	4	15.6	8	26.4
	Pennfield Townsh.	1	4.3	3	15.0	11	58.2
	Michigan State	1589	9.1	2936	15.3	8332	40.3

* = significantly higher than the Michigan rate, $p < 0.025$ one-sided

** = significantly higher than the Michigan rate, $p < 0.005$ one-sided

D.3 Hospital Discharge Data

Information on hospital discharges was obtained from the 1982 and 1983 Michigan Inpatient Hospital Data Base (MIDB82 and MIDB83, respectively). The principal source of data for these data bases was individual patient discharge records supplied by all Michigan short-stay hospitals through the Michigan Health Data Corporation (MHDC) and the Commission on Professional and Hospital Activities (CPHA). In addition, MIDB83 contains individual patient discharge records supplied to CPHA and MHDC by hospitals in the contiguous areas of Ohio, Indiana and Wisconsin. These databases include approximately 1.5 million records each year. The data in MIDB82 cover essentially 100% of the non-newborn short-term hospitalizations in Michigan hospitals. MIDB83 data cover essentially 100 %of non-newborn short-term hospitalizations for Michigan residents and nonresidents hospitalized in Michigan short-term hospitals. For the purpose of this study, nonresidents have been excluded from the MIDB83 data base. Data provided by the hospitals were carefully checked and edited by CPHA. Over 350 checks were performed while processing each record in order to ensure the validity of the data received. The data set was further reviewed and verified by the staff of the Office of the State Registrar and Center for Health Statistics. Although individual discharge records were provided, no specific patients or hospitals can be identified from the records. The location of the patient's residence is identified only by the postal zip code, rendering the zip code area the smallest geographic entity for which aggregate data may be tabulated.

To determine the rate of discharge for hospitalized residents of each county, the hospital use data, compiled by zip code, were aggregated to the county level using a set of allocation factors derived from the 1980 census for zip code areas which are in two or more counties. Rates can not be calculated for cities, townships and villages, but can be prepared for geographic areas that include and surround cities or villages with a post office and corresponding zip code. Specifically, discharge rates have been calculated for the Dowagiac and the Battle Creek areas which include the cities and surrounding areas.

Population values for specific zip code areas were derived from the published 1980 USA census figures. Hospital discharge rates for 1982 and 1983 for zip code areas were calculated using 1980 population numbers. No information on population by race was available at the zip code level when the discharge rates presented in this report were calculated. The discharge rates for the Dowagiac Area and Battle Creek Area zip codes (Dowagiac: 49047, Battle Creek: 49015 and 49017) are, thus, for all races. The Dowagiac Area includes all of Dowagiac City and a part of the surrounding rural areas. The Battle Creek Area includes most of Battle Creek City and a portion of the surrounding rural areas, which includes Springfield and the Verona area. It is estimated that the proportion of blacks in the Dowagiac and Battle Creek Areas are 9 and 10 %, respectively.

D.3.1 Results

Discharge rates for males are presented in Tables D.4.1 (males 1982), D.4.2 (white males 1983), D.5.1 (females 1982), and D.5.2 (white females 1983). Discharge rates for the two city areas and the two counties were compared to the Michigan rate using the procedure of Bailar and Ederer (1). No excess rates were observed in Cass county for either gender or year. The three other areas showed a striking similarity in significantly increased 1982 rates for diabetes, leukemia (males), and all-cancer (not statistically significant for Dowagiac females). The significantly increased discharge rate for liver disease for males in Battle Creek and Calhoun County (1982, 1983) and females (Battle Creek 1982 only) concurs with the increased mortality rate found in Battle Creek City in the period 1980-84. Fitting the overall pattern is the observation that pancreatitis, often associated with abnormal liver function and alcohol abuse, was more frequent in both sexes and years in Calhoun, Battle Creek and Dowagiac, although not

always statistically significant. As in the mortality data, increased rates for other diseases were present as well in an apparently haphazard fashion.

In evaluating these data, it should be kept in mind that discharge rates are not more than a very crude approximation of incidence rates. Since patients may be admitted in a hospital more than once for the same disease, discharge rates tend to be higher than incidence rates for not rapidly fatal diseases. Other diseases, such as diabetes, may not require regular hospitalization, and the annual discharge rate may therefore be entirely unrelated to the incidence rate. In general, however, trends in incidence and mortality rates may to some degree reasonably parallel discharge rate. However, annual fluctuations in small populations (county size or smaller) may be large enough to obscure existing real differences.

Rates for 5 year periods would have been more appropriate, but at the time this study was conducted, no more data were available than for 1982 and 1983. Yet, the most prominent findings from the discharge data show a striking similarity with those from the mortality study. Inconsistencies can be explained by the annual character of the discharge data, as compared to the 5-year period for mortality data, and the inability to break down discharge data to areas smaller than zip codes to distinguish the townships from Battle Creek City.

D.4 Summary

It can be concluded that both mortality and discharge data confirm the presence of increased rates for all cancer, leukemia, diabetes, liver diseases, and alcohol-dependency syndrome for some of the study areas, and for one or more of the periods studied. An increased discharge rate for pancreatitis fits in this pattern, but this disease was not included in the mortality study. Where the data base allowed separation of Battle Creek City from the adjacent Emmett and Pennfield townships, the areas containing the contaminated Verona area, the mortality rates were mostly lower than in Battle Creek City. The VOC levels were higher in Dowagiac than in the Verona area, and mortality rates were often higher than in the townships, but there was no consistent pattern. Moreover, the Dowagiac discharge and mortality rates were usually lower than the Battle Creek rates.

In summary, recognizing the insufficiencies of these kinds of data, the analysis of hospital discharge data and corrected mortality data has not resulted in evidence that VOC-contaminated water had caused observable excess morbidity or mortality rates. To the contrary, the rates for areas encompassing the exposed subpopulations tend to be lower than for other areas. However, since neither the mortality nor the discharge information could be analyzed in relation to individual exposure data, this study of rates for populations very much larger than the target populations in the Battle Creek health study should not be viewed as evidence against possible health effects of VOCs.

Table D.4.1 Hospital discharge rate per 1000 population by primary diagnosis for White Residents (State, Calhoun and Cass Counties) or Residents all Races (Battle Creek Area and Dowagiac Area). Males 1983.

Primary Discharge Diagnosis	Battle Creek Area	Calhoun County	Dowagiac Area	Cass County	Michigan State
All malignant neoplasms	8.961**	9.078**	11.581**	5.670	7.932
Malignant neoplasms of lower intestinal tract	0.686	0.895	1.616	0.915	1.037
Malignant neoplasm of the liver	0	0	0	0	0.044
Malignant neoplasms of bladder and kidney	0.553	0.530	1.616*	0.732	0.859
Leukemias	0.575**	0.514*	0.943**	0.320	0.323
Diabetes mellitus	2.522**	2.750**	4.713**	1.555	1.914
Alcohol dependency syndrome	4.359**	2.949**	1.077	0.686	2.470
Liver necrosis	0	0	0	0	0.022
Liver diseases	0.531*	0.480*	0.135	0.046	0.340
Pancreatitis	0.929**	0.480	1.077*	0.229	0.427
Kidney diseases	0.487	0.348	0.135	0.091	0.479

* Significantly higher than the Michigan rate, $p < 0.05$ one-sided

** Significantly higher than the Michigan rate, $p < 0.01$ one-sided

Table D.4.2 Hospital discharge rate per 1000 population by primary diagnosis for White Residents (State, Calhoun and Cass Counties) or Residents all Races (Battle Creek Area and Dowagiac Area). Males 1983.

Primary Discharge Diagnosis	Battle Creek Area	Calhoun County	Dowagiac Area	Cass County	Michigan State
All malignant neoplasms	7.678	7.790	10.504*	5.208	8.061
Malignant neoplasms of the esophagus	0.044	0.033	0	0.046	0.117
Malignant neoplasms of the stomach	0.376	0.251	0 *	0.046	0.163
Malignant neoplasms of the colon	0.708*	0.619	0.539*	0.372	0.484
Malignant neoplasm of rectum & rectosigmoid	0.133	0.134	0.943*	0.511	0.314
Malignant neoplasm of the liver	0.044	0.067	0.135	0.046	0.048
Malignant neoplasms of the bladder	0.509	0.485	0.808	0.372	0.667
Malignant neoplasms of the kidney	0.022	0.084	0.269	0.093	0.157
Leukemias	0.332	0.234	0.943**	0.418	0.305
Diabetes mellitus	1.770	1.985	3.232**	0.976	1.895
Alcohol dependency syndrome	4.226**	3.263	1.347	0.651	2.336
Liver necrosis	0	0	0	0	0.020
Liver diseases	1.173**	0.935*	0.808	0.372	0.673
Pancreatitis	0.907**	0.619*	0.808	0.372	0.440
Kidney diseases	0.620	0.519	0.943	0.372	0.474

* Significantly higher than the Michigan rate, $p < 0.05$ one-sided

** Significantly higher than the Michigan rate, $p < 0.01$ one-sided

Table D.5.1 Hospital discharge rate per 1000 population by primary diagnosis for White Residents (State, Calhoun and Cass Counties) or Residents all Races (Battle Creek Area and Dowagiac Area). Females 1982.

Primary Discharge Diagnosis	Battle Creek Area	Calhoun County	Dowagiac Area	Cass County	Michigan State
All malignant neoplasms	8.334*	8.588**	6.613*	3.895	7.513
Malignant neoplasms of lower intestinal tract	0.843	0.812	1.248	0.537	0.841
Malignant neoplasms of the liver	0	0	0	0	0.034
Malignant neoplasms of bladder and kidney	0.281	0.250	0.425	0.090	0.292
Leukemias	0.281	0.265	0.125	0.090	0.218
Diabetes mellitus	4.197**	3.763**	6.489**	2.865	2.850
Alcohol dependency syndrome	0.984**	0.671	0.374	0.224	0.649
Liver necrosis	0	0	0	0	0.016
Liver diseases	0.442*	0.297	0.374	0.179	0.248
Pancreatitis	0.783**	0.609**	0.499	0.269	0.382
Kidney diseases	0.723**	0.703**	0.624	0.313	0.376

- * Significantly higher than the Michigan rate, $p < 0.05$ one-sided
- ** Significantly higher than the Michigan rate, $p < 0.01$ one-sided

Table D.5.2 Hospital discharge rate per 1000 population by primary diagnosis for White Residents (State, Calhoun and Cass Counties) or Residents all Races (Battle Creek Area and Dowagiac Area). Females 1983.

Primary Discharge Diagnosis	Battle Creek Area	Calhoun County	Dowagiac Area	Cass County	Michigan State
All malignant neoplasms	7.310	7.901*	7.612	3.506	7.276
Malignant neoplasms of the esophagus	0.080	0.032	0	0	0.048
Malignant neoplasms of the stomach	0.301**	0.142	0.250	0.091	0.093
Malignant neoplasms of the colon	0.361	0.410	0.499	0.182	0.446
Malignant neoplasms of rectum & rectosigmoid	0.181	0.331	0.125	0.091	0.246
Malignant neoplasms of the liver	0.080	0.095*	0	0	0.038
Malignant neoplasms of the bladder	0.181	0.189	0.374	0.228	0.221
Malignant neoplasms of the kidney	0.161	0.095	0.125	0.046	0.092
Leukemias	0.060	0.047	0.374	0.228	0.201
Diabetes mellitus	3.394*	2.918	5.490**	2.459	2.756
Alcohol dependency syndrome	1.506**	1.088**	0.125	0.137	0.658
Liver necrosis	0.020	0.016	0	0	0.024
Liver diseases	0.502	0.473	0.624	0.273	0.530
Pancreatitis	0.562*	0.789**	0.250	0.091	0.378
Kidney diseases	0.502*	0.410	0.998*	0.364	0.384

* Significantly higher than the Michigan rate, $p < 0.05$ one-sided

** Significantly higher than the Michigan rate, $p < 0.01$ one-sided

Table D.6 Hospital discharge rates for 1982 and 1983 per 1000 population by primary diagnosis for Michigan Black and White residents. Males and Females.

Primary Diagnosis	1982		1983	
	Black	White	Black	White
All malignant neoplasms	7.718	7.056	7.660	7.143
Malignant neoplasms of the esophagus			0.082	0.230
Malignant neoplasms of the stomach			0.127	0.217
Malignant neoplasms lower intestinal tract	0.937	1.008	0.744	0.589
Malignant neoplasms of the liver	0.039	0.042	0.043	0.071
Malignant neoplasms of bladder and kidney	0.570	0.264	0.563	0.284
Leukemias	0.269	0.257	0.252	0.253
Diabetes mellitus	2.392	5.681	2.335	5.701
Alcohol dependency syndrome	1.540	3.148	1.480	2.795
Liver necrosis	0.019	0.031	0.022	0.028
Liver diseases	0.293	0.385	0.660	1.102
Pancreatitis	0.404	2.082	0.408	2.246
Kidney diseases	0.428	0.938	0.428	0.938

D.5 Reference

1. Bailar, J. C. and Ederer F. Significance factors for the ratio of a Poisson variable to its expectation. *Biometrics* 20: 639-643 (1964).

Appendix E

**Diabetes Mellitus and Exposure to Chlorinated
Volatile Organic Chemicals (VOCs) in Drinking Water**

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

E.1 Introduction

In 1981, the aquifer at the Verona Well Field, Battle Creek, Calhoun County, Michigan, was found to be contaminated with chlorinated, short-chain volatile organic chemicals (VOCs). Most of the city wells, and a large number of residential wells were also affected. Subsequently, VOCs were also detected in the water of a number of private wells in Springfield, which is adjacent to Battle Creek, and in Dowagiac, Cass County, Michigan. In 1984, Phase I of a two-phase epidemiologic study was initiated to collect demographic data to form exposed and reference cohort populations for Phase II, a retrospective followup study which started in July 1985. In Phase II, information was collected on morbidity and exposure for exposed and reference cohorts. This study involved 251 residents from exposed areas and 498 from comparison neighborhoods. The groups were matched for sex, age, family size, and neighborhood. Since 94% of the study population were white, race was not considered in the analysis.

The principles of risk assessment for chemicals were applied since the study focused on the potential effects of an earlier exposure. These included:

- 1) Exposure must be quantified as a dose and in time.
- 2) A dose-response effect must be determined.
- 3) A disease is attributable to exposure only if the exposure preceded the disease.
- 4) Quality control procedures must be extended to the quality of health data.
- 5) Uncertainties in the data and methods, and their effect on the study results, must be properly addressed.

Initially, there was no toxicologic or epidemiologic indication that diabetes mellitus was associated with exposure to VOCs. However, a pilot study of State-generated mortality rates for the period 1976-1980 by the Michigan Department of Public Health showed that the rate of diabetes mellitus was increased in Battle Creek City and Calhoun County compared to the overall State rates. Furthermore, a preliminary analysis of the Phase II data showed that more potential cases of diabetes were found in the exposed than in the comparison cohort. Although the number of these potential cases was small, an in-depth analysis of the association with exposure to VOCs appeared to be indicated.

E.2 Health Assessment and Case Definition

All hospitals serving the study area (Appendix B) were canvassed in a survey for the medical records for all study participants. In addition, physicians and hospitals specifically mentioned by the study participants were contacted to retrieve records. Copies of all death certificates were obtained. The records compiled were searched for information on diseases occurring as of 1970.

An in-person interview was conducted, some physical measurements were taken, and blood and urine specimens were collected from (fasting) participants 5 years and older. The clinical laboratory studies relevant to diabetes included serum and urinary glucose, and hemoglobin A₁C (HbA₁C). No provision was made for an oral glucose tolerance test.

The questionnaire was designed to maximize the recall of respondents by repeating some questions in different formats. (See Appendix A). The questions relevant to diabetes included: current disease or illness, medication, presence of diabetes, diabetes known prior to 1970, insulin

use, allergies (allergy to insulin was found), family members with diabetes, and any disease or condition not mentioned earlier. Each question also asked for the date of diagnosis, and name and address of hospital and physicians. Names of drugs and dates of prescription were recorded, and the drugs were screened for glucose-controlling pharmaceuticals (DIABMED).

This multisource approach to collecting data on the individual health status was designed to detect and correct recall failure, erroneous responses, and poor records. At this point, without any prior knowledge of exposure to VOCs, people were categorized according to the likelihood that diabetes was present, as described below.

DIAB=1 :

Any use of DIABMED (all persons using DIABMED had also a positive interview response and/or a positive medical record).

or

A positive response and/or medical record combined with either of the following: serum glucose and/or HbA1C which exceeds the upper normal limit of 110 mg% and 8.4% respectively, or a positive urinary glucose.

DIAB=2 :

A positive response and record, but no DIABMED used, no abnormal laboratory values. This category fits mild diabetes treated only with dietary restrictions.

DIAB=3 :

A negative response and record, no DIABMED, but glucose *and* HbA1C have values higher than 110% of the normal upper limit. People in this category can be considered borderline cases of diabetes.

DIAB=4 :

A negative response and record, no DIABMED, serum glucose higher than 160 mg%. Although incomplete fasting cannot be ruled out, at least one hour had elapsed between any possible food intake and blood sampling. Under these conditions a single glucose value of over 160 mg% is very suggestive of diabetes (1).

DIAB=5 :

A negative response and record, no DIABMED, serum glucose and HbA1C are marginally increased to maximally 110% of the upper limit.

DIAB=6 :

A positive interview *or* medical record, no DIAMED, no abnormal laboratory values.

A diagnosis was considered eligible for analysis if DIAB = 1 to 4. DIAB=5 and DIAB=6 were rejected because marginally elevated serum values are no evidence of diabetes in the absence of DIABMED use, a positive medical record, or the individual's awareness of having diabetes. In addition diabetes mentioned by one source only, either the respondent or a medical record, is not likely a case of valid diabetes, if no DIABMED is used and if the laboratory values are normal. False-positive responses may arise from misunderstanding information communicated by attending physicians, or misinterpreting an abnormal clinical test outcome (e.g., in pregnancy) as diabetes. False-positive medical records may be caused by erroneous entries into the medical file, a diagnosis based on a dubious laboratory test, or failure to update files, e.g., a one-time marginally abnormal test result unconfirmed in a followup medical evaluation. It should be kept in mind that the abstractors looked for the occurrence of a disease, not its absence.

To be eligible for the exposure-effect analysis (CASE=1), cases were required to belong in category HT=1-4, with the date of first diagnosis (TDIAG) later than 1-1-1970 *and* later than the date of moving to the study area (TIN). The people with DIAB = 5 or 6, and all others were

considered noncases. TDIAG was the earliest of: the date reported by the respondent; the date of the medical record; the clinical observation in case of DIAB=3 or DIAB=4; or the DIABMED prescription. In the analysis, TDIAG was corrected for a lag time of 1 year, that is, moved back one year as described in Chapter 6.

E.3 Exposure Assessment

The individual "total accumulated exposure" (TAE) level of study participants accumulated during the period of residency in the study area was estimated as the area under the $\log C = T$ curve, as described in Appendix C and elsewhere (2). In this notation, $\log C$ is the natural log of the concentration C of a given chemical in the water sample of a residential well (in parts per billion or ppb), and T is the time of sampling, expressed in months since 1-1-1970. The variable TAEVOC is the sum of the TAEs for the individual VOCs, expressed in units of ppb-months, and corrected for the TIN and the date a person stopped drinking the contaminated water (TSTOP). For those with diabetes, TAEVOC was calculated until the date of diagnosis (TDIAG) if this date was earlier than TSTOP.

In the analysis, several different exposure expressions were used. Exposure is defined as 1) having lived in the contaminated area, 2) TAEVOC greater than zero or not, or 3) TAEVOC with categorical (zero, low, and high) or continuous values. The border between *low* and *high* categories was the rounded median of all values of $TAEVOC > 0$. In either case, TAEVOC may include the seven VOCs of direct interest (TAEVOC proper), or the VOCs plus chloroform (Chl) from chlorinated city water. If chloroform was included, TAEVCL was used instead of TAEVOC. Finally, exposure was also expressed as the variables DOSEVOC and DOSEVCL, the product of WATER (volume of unheated tap water consumed at home per day) and TAEVOC and TAEVCL, respectively. As indicated in Chapter 6, these options offer opportunities for analyzing the exposure-effect associations from different viewpoints. However, DOSEVCL is the option most consistent with toxicology and risk assessment principles.

E.4 Statistical Analysis

The start of the followup period (TFU1) was 1-1-1970, or TIN if a person moved to the area after that date. The close of the followup period (TFU2) was 1) the date of diagnosis for cases, and 2) the date of death or the date of interview for non-cases. The total followup time $TFU = TFU2 - TFU1$. The analysis was done 1) as a crude analysis involving fourfold tables and rates of occurrence, and 2) as a multivariable analysis using Cox's proportional hazard model, worked out by Harrell as a SAS program named PROC PHGLM (3). This model is intended for classic cohort studies where the exposure level does not change over time from TFU1 to TFU2. In the Battle Creek scenario, as in most scenarios of environmental contamination, exposure levels were not constant over time, and people were not all exposed at the same time and for the same duration.

To solve this problem, PROC PHGLM was modified as described in Chapter 6. This version deals effectively with the fact that people may enter the cohorts at different points in time, that most exposed people had initially been unexposed, that exposure may stop before TFU2, and that exposure to VOCs may be followed by exposure to chloroform (Chl) from city water. In addition the computer program deals with *real* calendar time. As the number of cases ultimately eligible for analysis was too small for a model involving more than a few variables, the predictive variables were limited to the most essential ones; WATER (in models with TAEVOC or TAEVCL), WASH (a measure of water use for bathing or showering, see Equation 7, Chapter 4), and age at TFU1.

E.5 Results

During the investigation of state-wide mortality rates, it became obvious that many errors had been made in coding the residence of decedents and that the presence of diseases other than the cause of death had not been handled consistently. After making corrections, differences in diabetes mortality rates since 1970 between the State as a whole and Battle Creek City, Dowagiac, Calhoun or Cass Counties became statistically insignificant. For some of the above study areas, hospital discharge rates for diabetes in 1982 and 1983 were found to be increased when calculated by gender but not for the sexes combined. More detail is given in Appendix D.

Table E.1 Potential cases of diabetes and cases eligible for analysis in exposed (Expos) and reference (Refer) cohorts.

Category of Diagnosis	Total number DIAB > 0		Ineligible cases TDIAG < 1970		Ineligible cases TDIAG < TIN *		Cases eligible for analysis	
	Expos	Refer	Expos	Refer	Expos	Refer	Expos	Refer
DIAB=1	10	14	0	4	2	1	8	9
DIAB=2	1	0	0		1	0	0	0
DIAB=3	1	0	0		0		1	0
DIAB=4	1	1	0		0		1	1
DIAB=5	0	2	considered noncases					
DIAB=6	6	7	considered noncases					
Totals	19	24	0	4	3	1	10	10

* To prevent double counting, cases diagnosed before 1970 and before TIN were counted in the column TDIAG < 1970 only.

The study resulted in data which allowed a direct estimation of disease incidence. Forty three persons had some indication that diabetes might be present. Table E.1 shows a breakdown of this number by the category of validity of the diagnosis, and by eligibility for analysis. In total, 20 cases were eligible to enter the analysis; the number may be smaller for certain analyses, because some had missing values for variables used in the analysis. Table E.2 depicts the frequency of reports of confirmed diabetes (DIAB=1-4) by source of information. Interviews were the source with the highest yield of cases. On the other hand, "interviews only" also yielded the highest number (N=9), of false positive reports, compared to four false positives if medical records were the only source. Adding laboratory results as a second source of information increased the number of positive findings up to 100% of all confirmed cases for the combination interview/clinic, but the number of false positives was still not reduced to zero. Total coverage cannot be expected, since a medical file may have failed to record diabetes, and since normal laboratory values do not rule out diabetes if DIABMED is used. In any event, the combination of all four sources is the most appropriate approach to: ensure the retrieval of all cases; minimize the number of false negatives and false positives; and improve the accuracy of TDIAG. It should be noted that a higher yield of cases may also imply a higher yield of false positives. Although the DIAB categories were intended for separate analyses, the rareness of DIAB=2-6 rendered such analyses meaningless. Hence, for the analysis, the categories DIAB = 1-4 were pooled.

For the population-based incidence rate (IR), all cases of DIAB=1-4 diagnosed as of 1970 were counted (24 cases), regardless of whether the diagnosis was made before or after TIN. The

IR is the number of cases divided by the total TFU of 125,159 months (a correction factor of 12 is necessary to arrive at an *annual* IR). The IR for the period 1970-1985 for combined men and

Table E.2 Information sources for confirmed cases of diabetes.

Source	Number of confirmed cases
Interview only	25
Medical records only	21
Clinical examination only (DIAB = 3 or 4)	3
Clinical examination and/or interview	28
Clinical examination and/or medical records	24
Interview and/or medical records	25
DIABMED use and/or interview	25
DIABMED use and/or medical records	24
All four sources	28

women is: $24 \times 12 \times 1000 / 125,159 = 2.30$ per 1000 population/year (95% Poisson confidence limits of 1.47-3.42). The prevalence rate as of July 1985 is the number of all valid cases (ignoring TDIAG) over the total number of people in the study, *deleting those who were not alive as of July 1985*, and is equal to $25/736$ or 3.4 % (95% Poisson confidence limits 2.2-5.0). The prevalence and incidence rates for men are higher than those for women, although it is not statistically significant, as shown below.

Table E.3 Prevalence and incidence rate for diabetes with 95% confidence limits.

Prevalence (all ages)		Incidence rate/1000 all ages/yr	
males	4.7% (2.7-7.6)	males	3.34 (1.60-4.90)
females	2.3% (1.0-4.3)	females	1.60 (0.85-3.26)

Table E.4 depicts the distribution of cases and noncases by exposure status. The measures of association are the odds ratio (OR) and the relative risk or rate ratio (RR), calculated as described in Chapter 6. Cases were eligible for analysis if the (corrected) TDIAG was later than TIN, *and* later than 1-1-1970. Because exposure expressed as having lived in the VOC contaminated area is not related to when VOC exposure began, the eligibility criterion of TDIAG later than 1-1-1970 is irrelevant; hence the only eligibility criterion applied to this version of exposure was that the *uncorrected* TDIAG be greater than TIN.

Table E.4 indicates a statistically significant excess of diabetes among the exposed cohort, if exposure is defined as AREA, or as TAEVOC and DOSEVOC. The dose-response effect is negative (higher risk at lower TAEVOC/DOSEVOC). The OR and RR for TAEVCL and DOSEVCL are below unity (excess in the reference cohort) or slightly above, and are statistically insignificant. The results remain essentially unchanged if the analysis is limited to the Verona exposed and Calhoun county reference cohorts. The negative direction of the dose-response effect resulted in not a single case among people of the Verona area exposed to the higher levels of DOSEVOC and DOSEVCL.

Table E.4 ORs and RRs for diabetes (DIAB) and exposure expressed as a dichotomous (yes or no resident of exposed area, yes or no positive value for TAE and DOSE variables) and categorical variable (zero-low-high).

Total Population 1970 - 1985: Exposure expression is AREA or TAE														
Diabetes	Contam. AREA		Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)								
	No	Yes	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed No	Exposed High	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed No	Exposed High
	No	483	238	511	210	511	105	511	105	236	485	236	375	236
Yes	11	12	11	9	11	5	11	4	8	12	8	8	8	4
OR =	2.21		1.99		2.21		1.77		0.73		0.63		1.07	
p =	0.05		0.10		0.13		0.25		0.32		0.25		0.57	
RR =			2.06		2.47		1.71		0.86		0.76		1.16	
p =			0.09		0.09		0.26		0.45		0.38		0.51	

Total Population 1970 - 1985: Exposure expression is DOSE													
Diabetes	Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)									
	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed No	Exposed High	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed No	Exposed High	
No	526	182	526	90	526	92	266	442	266	345	266	97	
Yes	10	9	10	6	10	3	7	12	7	9	7	3	
OR =			2.60		3.51		1.72		1.04		0.99		1.19
p =			0.04		0.02		0.31		0.58		0.59		0.53
RR =			2.77		3.90		1.75		1.21		1.18		1.31
p =			0.03		0.01		0.30		0.44		0.47		0.47

continued on next page

Table E.4 - Continued.

Verona Exposed + Calhoun Reference 1977 - 1985: Exposure expression is TAE.									
Diabetes	Contam. AREA		Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)			
	No	Yes	Exposed No	Exposed Yes	Exposed No	Exposed Yes	Exposed No	Exposed Yes	
			Low	High	Low	High	Low	High	
No	386	172	409	148	409	90	418	360	58
Yes	7	8	5	5	5	4	3	6	1
OR =	2.56		2.76	3.64	1.41		0.78	0.77	0.80
p =	0.06		0.10	0.07	0.55		0.48	0.48	0.66
R =			2.91	3.78	1.51		0.86	0.85	0.91
p =			0.09	0.06	0.53		0.53	0.54	0.71
Verona Exposed + Calhoun Reference 1977 - 1985: Exposure expression is DOSE.									
Diabetes			Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)			
	No	Yes	Exposed No	Exposed Yes	Exposed No	Exposed Yes	Exposed No	Exposed Yes	
			Low	High	Low	High	Low	High	
No			420	125	420	74	380	324	56
Yes			5	5	5	5	3	7	0
OR =			3.36	5.68	0		1.01	1.99	0
p =			0.06	0.01	0.57		0.63	0.55	0.42
RR =			3.63	5.87	0		1.12	1.29	0
p =			0.05	0.01	0.61		0.59	0.50	0.48

- 1) p-values were estimated by a one-sided Fisher's exact test.
- 2) The total number of people and cases in the DOSE tables are somewhat smaller than the total number in the TAEVOC tables, as people with missing data for WATER (and thus for DOSE) were deleted.
- 3) The threshold between low and high exposure level is determined by the median of all non-zero values. Low and high are relative terms for the analysis only. They do not have a toxicological connotation.
- 4) Since the only eligibility criteria for exposure expression "contam. area" is that the uncorrected TDIAG > TIN, the number of cases in the OR table for "contaminated area" is larger than in the other fourfold tables. The OR and RR for this exposure are presented for the sole purpose of demonstrating the effect on these ratios of a conventional case and exposure definition.

Table E.5 Results of proportional hazard analysis, blocking for period. The values shown refer to the coefficient (beta) and the p-value (p) of the variables in the model. AGE1 is the age at the start of the followup, N is the number of cases with no missing values for the variables in the model, and R is the R-statistic akin to the multiple correlation coefficient in the normal setting (3).

Cohorts involved	Model 1 (N=17, R=0.219)			Model 2 (N=17, R=0.225)		
	Variable	beta	p	Variable	beta	p
All cohorts	LTAEVCL	0.0312	0.6206	LDOSEVCL	0.0610	0.2930
	WATER	- 0.0967	0.2235			
	WASH	- 0.0160	0.5842	WASHVCL	- 0.0000	0.5970
	AGE1	0.0503	0.0007	AGE1	0.0498	0.0003
Verona exposed and Calhoun reference	Model 3 (N=9, R=0.159)			Model 4 (N=9, R=0.201)		
	LTAEVCL	0.0593	0.5846	LDOSEVCL	0.0614	0.4796
	WATER	- 0.0477	0.6251			
	WASH	0.0036	0.8805	WASHVCL	0.0000	0.8351
	AGE1	0.0565	0.0037	AGE1	0.0556	0.0030

In all models the R-statistic was entirely the effect of AGE1 (the partial R for all other variables is 0.000).

The results of the multivariable analysis for TAEVCL and DOSEVCL are shown in Table E.5. The exposure variables have been transformed into their natural logarithms (LTAEVCL and LDOSEVCL). In the model with DOSEVCL, WATER is deleted as it is already incorporated in the DOSE variable, and WASH is converted into the variable WASHVCL, equal to WASH x CS (CS = sum of concentrations of VOCs in the current water sample plus 6 ppb, which is the chloroform concentration for residents on city tap water). The models apply to exposure calculated as TAE and DOSE accumulated since TFU1 up to TFU2 (entire followup period).

The coefficients for the exposure variable were positive, in agreement with the ORs and RRs above unity in Table E.4, but the p-values for TAEVOC and DOSEVOC were much larger, and WATER and WASH were negatively associated with the risk of contracting diabetes. This argues against an hypothesis of a toxic effect of VOCs. The change of the sign of the coefficient of the WASH variable (if the analysis is restricted to the Verona exposed and the Calhoun County reference cohorts) probably reflects the unreliability of a multivariable analysis with only nine cases in the data set. Similar results were obtained from analyses with exposure incurred in the index period only, or exposure computed as TAE or DOSE accumulated from TFU1 until the starting date of the index period (Chapter 6). No association was found between serum glucose or HbA1C and TAE or DOSE in correlation tests and multilinear regression with sex, age, the body mass index, and blood lipids as covariates.

To identify risk factors for diabetes in the absence of variables related to exposure to VOCs (TAE, DOSE, WASH, WATER), the data were also analyzed with the conventional PROC PHGLM (no partitioning of the period of observation), while ignoring eligibility criteria with regard to TDIAG. Because of the different epidemiology of juvenile diabetes, the analysis was restricted to people older than 20 years of age, while deleting cases who were younger at the time that diabetes was first diagnosed. The basic model contained the regressors SEX and FAMILY history

of diabetes, the only risk factors present at TFU1. In this analysis, age was already controlled for, since everyone was of the same age at TFU1. With 26 cases and 463 noncases, the model R-statistic was as high as 0.21, $p(\text{SEX}) = 0.1024$, and $p(\text{FAMILY}) = 0.0008$. The OR (and its 95% confidence limits) derived from the regression equation, was 1.94 (0.88 - 4.30) for being a male, and 4.10 (1.80 - 9.34) for having a positive family history.

In an attempt to identify additional risk factors, other variables were tested as a third covariate. Caution is needed in interpreting the results of such an exploratory analysis, because these additional factors represent the status at TFU2 and not at TFU1. Further, the significance value may be misleadingly low because of making multiple comparisons. Additional risk factors with $p < 0.1$ were: the Quetelet Index or QI (equal to weight kg/height m^2), GGTP (liver enzyme), blood urea nitrogen, triglycerides, and hypertension.

The body mass index (QI) was quantitatively the most important risk factor. The model with SEX, FAMILY, QI (and AGE as this is built in into the model), had an R-statistic of 0.41, twice as high as without QI. Unrelated to diabetes were smoking, alcohol use, education, income, occupational exposure to chemical compounds (VOCs among others), hypertension drugs, cholesterol, SGPT-SGOT-LDH (liver enzymes), urinary beta2- microglobulin (kidney function test), creatinine, uric acid, along with DDT, PCB, and PBB. When adding a third covariate, it became clear that the OR for being a male was too low because of lack of control for confounding factors. For instance, the OR increased to 3.10 (95% limits 1.27-7.58) when controlling for QI.

E.6 Discussion

As a component of Phase II, the Michigan Department of Public Health analyzed State-generated mortality data to look at possible increased mortality rates in political subdivisions encompassing the study area. A small and statistically insignificant increase was found for the mortality rate of diabetes in Battle Creek City. As shown in tables D.4 and D.5 of Appendix D, the hospital discharge rates for diabetes was often significantly increased for Battle Creek City, Dowagiac, and Calhoun and Cass Counties. Discharge rates, although a better measure of disease incidence or prevalence than the mortality rate, are still a poor parameter of the risk of disease from toxicants. The discharge rate of a disease requiring regular medical attention, such as diabetes, heavily depends on factors governing local access to and utilization of medical facilities. The mortality and hospital discharge rates apply to political subdivisions that very poorly match the contaminated areas. The study of disease incidence in the affected population and a reference population is the proper approach to evaluating the effect of VOC exposure.

From a risk assessment viewpoint, the main problem in collecting health data is to ascertain the accuracy and the completeness of the data obtained. Multiple information sources were utilized to determine the current and past medical history. An analysis of discrepancies between the various data sources had the following results. Out of four people with a positive interview, confirmed by laboratory data but not by a medical record, three had not been hospitalized, and the fourth was hospitalized prior to 1970. In the last case, there could have been a hospital record, but the search for records did not cover the period before 1970. With regard to the first three cases, a failure to retrieve records from private physicians' offices is a possibility, assuming that the records were complete. Eight people reported diabetes but had no DIABMED, no positive medical record, and normal laboratory values. Five of these said they had not been hospitalized. One had been hospitalized prior to 1970, one since 1970 (out-of-state, diagnosed before TIN, no hospital record found), and one did not know. These are apparently false-positive interview responses, possibly by misinterpretation of medical information received from private physicians. Harris *et al* found that 14% of self-reported diagnosis may be false positive (4).

A positive medical record was found for five persons with a negative interview response and no DIABMED. One of these five interviews was made with a next-of-kin of a subject who died of cardiovascular disease, and no clinical laboratory studies were available. No diabetes was reported on the death certificate or in any of the many other medical records. The positive medical record is, therefore, probably based on a single abnormal glucose value or dubious OGTT, which was unconfirmed by subsequent followup. "Chemical diabetes" was a physician's diagnosis of the second case, indicating dubious or borderline diabetes based on either a single glucose measurement or a single impaired OGTT (5, 6), and not diabetes. Other medical records did not reveal diabetes or another abnormal OGTT. The other three persons were a father and his two sons. Both children had a hospital record of "history of diabetes" made on the very same date, but there was no information on the date of diagnosis or confirmatory blood tests. None of these persons or their family members reported a family history of diabetes. These cases hence suggest an erroneous entry in the medical record or an erroneous source of these entries.

In summary, there were probably eight false positive interview responses (8/33 or 24% of all positive responses) and five false-positive medical records (5/26 or 19% of all positive records). This led to disregarding one-source-only cases of diabetes for analysis (category DIAB=6) except highly abnormal laboratory tests, to reduce the possibility of misclassification. To retrieve medical records, the contractor's investigators visited all hospitals serving the study area. The out-of-the-area hospitals and physicians with a private office, if named by an interviewee, were requested to mail a copy of the full record. As record retrieval in these cases was done by local staff, it is conceivable that some files could have ended up incomplete. The data suggest that there might have been a failure to retrieve up to four medical records of diabetes (1 in the exposed and three in the reference cohort), confirmed by diabetes therapy and abnormal glucose levels, and possibly all from private physician's offices. It may be concluded that the results of the search was quite effective for hospital records, but less so for physicians' records.

Of the 21 cases with a positive interview response *and* a medical record, three had a missing date of diagnosis in either the interview response or the record, and four had no difference in these dates. The date reported in the interview was earlier than the medical record date in six cases (range from 18 to 129 months, average 52 months), and later in eight cases (range from 3 to 120 months, average 35 months). For the analysis, the earliest of the recorded dates (including the dates of drug prescription and clinical examination) was selected as the date of diagnosis. Although many studies have identified the problem of "recall deficiency", no study was found that examined the validity of the date of diagnosis mentioned during an interview. Out of 18 subjects with dates from both interview and medical record, eight mentioned a date up to 10 years later than the recorded date of hospital discharge. The potential for so large an error should be seriously considered when conducting an investigation into the relation of a disease and a predefined exposure. The discrepancies and deficiencies were unrelated to gender, age, or exposure status, and are therefore unlikely to have biased the results of the exposure-effect analysis.

A simple check on the completeness of case-finding is to compare the crude IR and the prevalence. The annual IR of 2.3/1000 for 1970-1985 is slightly lower than the 2.7/1000 estimated for the 1978 U.S. population by Herman *et al* (7) from information on the prevalence compiled by the National Center for Health Statistics (NCHS) in the National Health Interview Survey (NHIS). However, NHIS data do not permit estimating an IR; in the NHIS, people were asked about diabetes in the past 12 months, not about the *first* time that diabetes was diagnosed, and there was also no validation of the responses. As shown here, TDIAG from interviews was quite unreliable, and often a self-reported diagnosis was not confirmed by other data. The IR for other populations in the U.S. is comparable to that observed in the Battle Creek Study (except for much higher IRs for Indian tribes), as shown in a series of incidence studies reviewed by Everhart *et al* (8). In contrast, the *prevalence* of 2.3% in the U.S. measured in the NHIS (9), is only two thirds of the 3.4% observed in the current study.

As the NHIS solely relies on interviews, the true prevalence is probably better reflected in the results from another NCHS investigation: the National Health and Nutrition Examination Survey 1976-1980 (NHANES), based on interview and an OGTT. The prevalence estimated from interviews was 2.8% for males and 3.6% for females for the age-group 20-74 years (10). By including all others who were diabetic by their OGTT, using National Diabetes Data Group criteria (6), the NHANES rate increased to 5.3% and 7.0%. The rates in the current study for age 20-74 are 7.0% (95% confidence limits 3.9-11.6) for males and 2.7% (1.1-5.5) for females. That is, the prevalence among females is half of the NHANES rate, although the difference is insignificant at $p > 0.05$. No explanation was found for the lower prevalence among women, which resulted in rates higher for men than for women, the reverse of what was found in all other studies.

By accepting category DIAB=6 as valid cases, the prevalence would be similar or higher than the total (response and laboratory cases) NHANES level (males 8.5%, females 4.9%). Although some people with DIAB=6 could have been true cases of diabetes, the occurrence of false-positive responses or records must not be ignored in an exposure-effect analysis. It is common in medical practice to consider a person with a single abnormal laboratory finding as a potential case of diabetes until the results of medical followup are known. This practice could easily have caused a false-positive response or record in the current study, as no search was made for negative outcomes of a medical followup. The NHANES criteria for a diabetic OGTT were more rigid than those usually in place in general medical practice. A dubious OGTT does not constitute diabetes, and the risk of diabetes developing from an abnormal OGTT is not well known (5, 9). It would have been most informative if, in the NHANES, people who said they had diabetes had also been subjected to an OGTT. Misdiagnosis, especially if the diagnosis was based only on an OGTT, is a possibility not to be overlooked as shown by a review of the NHANES study (4), and by the findings in the current study. In summary, the overall impression is that the prevalence of diabetes in the study population is comparable to that in the U.S. population and, thus, there is no evidence of an incomplete retrieval of cases.

It has been shown (Table 4.3, Chapter 4) that 11% of the people who lived in the contaminated areas had not been really exposed, which illustrates the danger of conducting a statistical analysis with an improper expression of exposure and no ascertainment of the temporal sequence of exposure and effect. Yet, this kind of analysis has been quite common in environmental studies. The effect of misclassification can be serious, as demonstrated in Table E.4.1. The ORs and RRs decrease considerably with an improved exposure expression. The one most consistent with risk assessment principles is DOSEVCL, for which the ratios are more or less equal to unity. The large ORs and RRs for TAEVOC and DOSEVOC are negated by the clearly negative dose-response effect.

The number of cases was too small for multivariable analysis, even for models with only three or four regressor variables, rendering the value of the outcomes shown in Table E.5 dubious. The positive coefficients for the exposure variables are accompanied by very large p-values, and the coefficient for WATER is negative. The analysis of the data from a population limited to the Verona exposed and the Calhoun County reference cohort has the advantage of better exposure data. Further limiting the cases to those diagnosed in the period 1977-1985, with the inherent advantage of more reliable and complete health data would have been the ideal approach, were it not for the associated substantial reduction in the number of cases. Ignoring the large p-values, Tables E.4 and E.5 show that the analysis of the restricted study population is in agreement with that of the total study population. If the Verona exposed population is judged on its own area (no reference area), the dose-response effect for all exposure expressions was negative, and not a single case occurred at the higher exposure levels. This finding, too few eligible cases for multivariable analysis including more covariates, and the negative dose-response effect shown in the ORs and RRs from fourfold tables, are sufficient to conclude that there is no evidence that VOCs played a role in the development of diabetes.

Other alternative analyses included stratification by sex, the acceptance of diagnostic categories DIAB=5 and 6, a case definition limited to DIAB=1, the sum of the current water sample concentrations of VOCs as an expression of exposure, and dropping the eligibility criterion of TDIAG later than 1-1-1970. None of these analyses had results significantly different from those described earlier. This consistency inevitably leads to the conclusion that there is no evidence of a toxic effect of VOCs at levels prevailing in the drinking water of the study area.

Reports on the association of hypertension and diabetes are conflicting, and suggestions have been made that discrepancies in reported results are related to sampling errors, or to different categories of diabetes and hypertension. Similar conflicting associations were reported for diabetes and blood pressure (5, 11, 12, 13). In the current study, a statistically significant association (OR=2.59 with $p=0.02$) was found for hypertension, controlling for age, and sex. However, there was no association if hypertensive cases were limited to those that used anti-hypertensive drugs, or if the body mass index was added to the model. Hence, the association of diabetes with hypertension might have been confounded by obesity. Studies reporting on an association with hypertension may have suffered from a lack of control for this factor. It has been postulated (11, 13) that hypertension, while present in most cases of noninsulin dependent diabetes mellitus (NIDDM), would be a late complication of insulin dependent diabetes mellitus (IDDM). In the current study, the data collected was insufficient to separate these types of diabetes with certainty.

The analysis of adult-onset diabetes (age at diagnosis older than 20 years) in a cohort analysis starting at age 21, with sex and family history as the only regressors, yielded an OR for males versus females of 1.94, which is somewhat lower than the male/female ratio of the prevalence and the incidence rates. The OR increased to about 3, however, when the model was controlled for confounding factors. Quantitatively the most important risk factors were a positive family history and high serum values for triglycerides. It should be emphasized that, with juvenile-onset diabetes excluded, the remaining cases were predominantly NIDDM cases. Given the high odds for people with a positive family history to develop diabetes, a more detailed investigation of the kind of diabetes and the role of the family history as a risk factor may be of public health importance.

In addition to an analysis of the nature of the statistical association between diabetes and known or suspected risk factors, this study yielded other findings of public health importance. The first is that out of 18 persons who knew they had diabetes, and who had clearly elevated serum glucose (range 121 - 228 mg%, in one case with HbA1C = 12%), five had no pharmacotherapy and four of these had a positive medical record. Five other individuals had abnormal laboratory values with no indication of diabetes from interview or medical record. Three of these had blood tests strongly suggestive of diabetes: two persons had a fasting glucose level of 169 and 173 mg% and normal HbA1C values; and the third had a glucose level of 126 and HbA1C of 9.8 (15 and 16%, respectively, above the upper limit of the normal range). The remaining two persons had marginally elevated (less than 10% above the upper limit) values for glucose and HbA1C, and were not included in the statistical analysis. Thus, at least eight out of 28 confirmed cases (29%) were people with untreated diabetes. Further, there were three (11%) newly found cases of diabetes, much less than the NCHS estimate of 50% undetected cases (7). Although some degree of elevated fasting glucose level can be tolerated presumably without risk, the health risk of leaving diabetes untreated is well known.

A third unexpected finding is that HbA1C, considered a parameter for monitoring the course of diabetes control better than glucose itself (11), did not prove to be a good indicator of the presence of diabetes. Of the people represented in Table E.2, 38 confirmed cases had laboratory values, 11 of these had abnormal glucose and normal HbA1C levels. Of these 11 cases, none used insulin, five were on oral pharmacotherapy, and six had no medication. Among the 689 people with complete laboratory records, three had a combination of normal glucose (range 80 - 98 mg%)

and a HbA1C value exceeding 110% of the upper normal limit of 8.4 (range 10.3 - 10.9). These three subjects had no indication of diabetes from the interview, medication, or medical record. It appears that HbA1C is not a suitable screening tool in the general population, while its value as a supporting tool for diabetes control remains to be established. Serum glucose alone would have been a quite satisfactory screening tool. Of the 23 people with confirmed diabetes, 20 had an abnormal glucose value, two had a normal value but were on medication, and one had a normal glucose (and HbA1C) level while not on medication. This is equal to a sensitivity of the glucose test of 0.87. The specificity of the test cannot be estimated from the data, however, until the five persons with abnormal glucose values and no positive record have been subjected to a followup examination to confirm or reject the presence of diabetes.

E.7 Conclusions and Summary

Following the detection of groundwater contaminated with chlorinated volatile organic compounds (VOCs), a retrospective cohort study was conducted involving 749 people with an average followup period of approximately 13 years, and a 1:2 ratio of exposed to unexposed subjects. The study design adhered to principles of risk assessment for assessing the health risk from chemicals. The study was not specifically designed to address diabetes mellitus, but the presence of 43 potential cases and a possible excess among residents of the contaminated areas justified a detailed analysis. After validation of the available information, 28 valid cases remained, equal to a prevalence rate of 3.3% as of July 1985, and an annual incidence rate for the period 1970-1985 of 3/1000 males and 1.8/1000 females of all ages. The incidence male/female ratio was higher than observed in samples of the U.S. population. After applying the eligibility criteria, 20 cases remained available for analyzing potential exposure-effect associations. The following conclusions can be drawn from this study:

- 1) Exposure to VOCs in drinking water is not related to the risk of developing diabetes.
- 2) Quantitatively the most important risk factors for diabetes are a positive family history and obesity (high body mass index). Other risk factors include being a male (which may have been a chance effect, as most studies found an excess of cases among females), age, and high serum values of triglycerides, GGTP (liver function test), and blood urea nitrogen.
- 3) DDT, PCB, and PBB (at the prevailing serum levels), education and income level, occupational exposure to a variety of chemicals (among others VOCs), smoking, alcohol, kidney function tests, liver function tests other than GGTP, and cholesterol have not shown an association with the occurrence of diabetes, but the number of cases did not allow an analysis that included concurrent diseases and different types of diabetes.

Other important findings from the study are: a) an historical survey for medical data limited to an interview and clinical examination or with a medical record survey as a second source of information is bound to underestimate the disease incidence and prevalence, and does not permit adequate control of the quality of the health data compiled; b) there is a substantial inaccuracy in the dates of diagnosis of diabetes reported by the interviewees; and c) hemoglobin A₁C is a poor indicator of the presence of diabetes if the test is used as the sole parameter.

E.8 Addendum

The incidence and prevalence of diabetes have been recalculated, using an expanded DIAB=4 diagnostic category, which also includes cases with a fasting blood glucose of 140-160 mg% with a negative interview and a negative medical record. This is to accommodate a WHO definition of diabetes solely based on a fasting glucose level. The incidence and prevalence rose by less than 10% (2 cases added); there was *no* change in the lack of association between diabetes and exposure to VOCs. A scientific publication on the new incidence rate is currently in preparation.

E.9 References

1. Forsham, P.H., Steinke, J., and Thorn, G. W. Diabetes mellitus. In: *Principles of Internal Medicine*. Harrison T.R., et al., eds., McGraw Hill Co., New York NY, 5th ed. 1966, p.493.
2. Freni, S. C. and Phillips, D. L. Estimation of the time component in the movement of chemicals in contaminated groundwater. *Environ. Health Perspect.* 74: 211-221 (1987).
3. Harrell, F. E., Jr. The PHGLM procedure. In: *SUGI Supplemental Users Guide*, SAS Institute Inc., Cary NC, 5th edition 1986, p. 438-466.
4. Harris, M. I., Hadden, W. C., Knowler, W. C., and Bennett, P. H. Prevalence of diabetes and impaired glucose tolerance and plasma glucose levels in the U.S. population aged 20-74 yr. *Diabetes* 36: 523-534 (1987).
5. World Health Organization Study Group. *Diabetes Mellitus; Technical Report Series 727*, World Health Organization, Geneva, 1985.
6. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 28: 1039-1057 (1979).
7. Herman, W.H., Sinnock, P., Brenner, E., et al. An epidemiologic model for diabetes mellitus: Incidence, prevalence, and mortality. *Diabetes Care* 7: 367-371 (1984).
8. Everhart, J., Knowler, W. C., and Bennett, P. H. Incidence and risk factors for noninsulin-dependent diabetes. In: *Diabetes in America*. National Diabetes Data Group, DHHS Publ. (NIH) 85-1468, 1985, Chapter IV, pp. 1-35.
9. National Center for Health Statistics. Prevalence of selected chronic conditions - United States, 1979-81. *Data from the National Health Survey Series 10, No. 155*. DHHS Publ. no. (PHS) 86-1583, Hyattsville, MD, 1986.
10. National Center for Health Statistics. Prevalence of diagnosed diabetes, undiagnosed diabetes, and impaired glucose intolerance in adults 20-74 years of age. United States, 1979-81. *Data from the National Health Survey Series 11, No. 237*. DHHS Publ. (PHS) 86-1583, Hyattsville, MD, 1987.
11. Pyorala, K., Laakso, M., and Uusitapa, M. Diabetes and atherosclerosis: an epidemiologic view. *Diabetes/Metabolism Reviews* 3: 463-524 (1987).
12. Jarrett, R. J., Keen, H., McCartney, P., et al. Glucose tolerance and blood pressure in two population samples: Their relation to diabetes mellitus and hypertension. *Int. J. Epidemiol.* 7: 15-24 (1978).
13. Ritz, W., Hasslacher C., Tschöpe W., et al. Hypertension in diabetes. *Contr. Nephrol.* 54: 77-85 (1987).
14. Bunn, H. F. Evaluation of glycosylated hemoglobin in diabetic patients. *Diabetes* 30: 613-617 (1981).

Appendix F

**Hypertension and Exposure to Chlorinated
Volatile Organic Chemicals in Drinking Water**

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

F.1 Introduction

In 1981, the aquifer at the Verona Well Field, Battle Creek, Calhoun County, Michigan, was found contaminated with chlorinated, short-chain volatile organic chemicals (VOCs). Most of the city wells, and a large number of residential wells were affected. Subsequently, VOCs were also detected in the water of a number of private wells in Springfield, adjacent to Battle Creek, and Dowagiac, Cass County, Michigan. In 1984, Phase I of a two-phase epidemiologic study was initiated, basically to collect demographic data to form exposed and reference study populations for the subsequent Phase II, a retrospective followup study, starting in July 1985. In Phase II, information was collected on morbidity and exposure for exposed and reference cohorts. The study involved 251 residents from contaminated areas and 498 from comparison neighborhoods, group matched for sex, age, family size, and neighborhood. As 94% of the study population were white, race will not be considered in the analysis.

Since the study focused on the potential effects of a predefined exposure, the principles of risk assessment for chemicals were applied. These included:

- 1) Exposure must be quantified as a dose and in time.
- 2) A dose-response effect must be determined.
- 3) A disease is attributable to exposure only if the exposure preceded the disease.
- 4) Quality control procedures must be extended to the quality of health data.
- 5) Uncertainties in the data and methods, and their effect on the study results, must be properly addressed.

There was no toxicologic or epidemiologic indication that hypertension was associated with exposure to VOCs. However, a superficial evaluation of disease frequencies resulting from the study showed a large number of possible cases of hypertension, and a further analysis of the association of this disease with VOC exposure and/or other risk factors seemed indicated.

F.2 Health Assessment and Case Definition

All hospitals serving the study area (Appendix B) were canvassed in a survey for medical records of all study participants. In addition, doctors' offices and all hospitals mentioned by study participants were contacted to retrieve records. Copies of all death certificates were obtained. These records were searched for information on diseases occurring as of 1970.

An in-person interview and a limited clinical examination were conducted. Blood and urine samples were collected from (fasting) participants aged 5 years and older. The blood pressure was measured in a sitting position. Three readings were made, with a pause of at least 2 minutes in between. The Hawksley random-zero sphygmomanometer was used. This eliminated the examiner's tendency to bias the second and third readings towards the first (1). Since an improper cuff size could lead to erroneous readings (2), four sizes were made available to the technician. The cuff size selection was determined by the circumference at the midpoint of the upper arm. Cuffs were deflated at a rate of approximately 2 mm mercury (Hg) per second. The diastolic pressure was the Korotkoff 5 phase, or Korotkoff 4 (rarely) if there was no fifth phase.

The questionnaire was designed to maximize the recall ability of respondents by repeating questions in different formats (Appendix A). The questions relevant to hypertension included: any current disorder, presence of hypertension, hypertension known prior to 1970, any prescribed or

over-the-counter medication, and any disease or condition not mentioned previously. Each disease-specific question also asked for dates of diagnosis, and names and addresses of hospitals and physicians. Names of drugs and dates of prescription were recorded, and the drugs were screened for those with a known blood pressure lowering effect (BPMED). At the visit to the clinic, people were asked about diseases that might have occurred since the interview, and questions were repeated about possible anti-hypertensive medication. This multisource approach to collecting data on the health status was designed to detect and correct recall failure, erroneous responses, and poor records.

The health data compiled were screened for any mention of hypertension, use of BPMED, or an average of the three blood pressure readings (made at the clinic) of 140+ mm Hg systolic (SP) or 90+ mm Hg diastolic blood pressure (DP). Out of 724 people eligible for the clinic (age 5 years and older), 35 had no pressure measured (13 deceased, 22 did not show up, of whom eight were children). Established hypertension was defined as $SP > = 160$ mm Hg and/or $DP > = 95$ mm Hg. Borderline hypertension existed if $139 < SP < 160$ and/or $89 < DP < 95$ mm. Without prior knowledge of the exposure status, people with any indication of possible hypertension were categorized according to the likelihood that hypertension (HT) was present, as described below.

HT = 1:

Use of BPMED in combination with a positive interview response and/or medical record, regardless of the blood pressure;

or

Clinical examination showed an established hypertension, regardless of use of BPMED, interview response, or medical record;

or

One special case with a positive medical record, but without data on BPMED or blood pressure was accepted in this category. It involved a person who died before the study began. The next-of-kin was unable to provide information, but paralysis following a cerebrovascular accident was listed on the death certificate, which fitted the picture of a case of serious hypertension with fatal outcome.

HT = 2:

Borderline hypertension with a positive interview response and/or medical record, but no BPMED.

HT = 3:

Positive response and/or record, but no BPMED, and blood pressure normal or unknown.

HT = 4:

BPMED is used, but negative response and/or record, and blood pressure normal or unknown.

A case was considered valid if HT = 1 or 2. There were no children under age 16 with hypertension according to the criteria for children (3). Categories HT=3 and HT=4 were rejected for the following reasons:

HT = 3:

A positive response and/or record unsupported by BPMED or elevated blood pressure is likely to be a false positive. Such a response may be caused by lack of understanding of medical information. A false positive medical record may be caused by erroneous file entries, or by a single finding of elevated blood pressure. If subsequent medical followup showed that the subject was normotensive, the abstractors would not have noted it, as they searched for diagnoses, not for the absence of disease. It cannot be ruled out, however, that some people in this category, who had unknown blood pressure, might have had elevated values.

HT=4:

Use of BP MED by itself is no evidence of hypertension, since the drugs may have been prescribed to treat disorders other than hypertension.

To be eligible for the exposure-effect analysis, cases were required to belong in category HT=1 or 2 (valid case) with the date of first diagnosis (TDIAG) later than 1-1-1970 and later than the date of moving to the study area (TIN). People with HT=3 or 4, and all others were not considered as cases. TDIAG was the earliest of: the date reported by the respondent; the date found in the medical record; the clinical observation; or of BP MED prescription. In the analysis, TDIAG was corrected for a lag time of 1 year.

F.3 Exposure Assessment

Individual exposure levels, accumulated during residency in the study area, were estimated as TAE (total accumulated exposure), equal to the area under the $\log C = T$ curve, described in Appendix C and elsewhere (4). In this notation, $\log C$ is the natural logarithm of a chemical compound's concentration C in well water (in parts per billion or ppb), and T is the date of sampling the water in months since 1-1-1970. The composite value TAEVOC is the sum of the specific chemical TAEs, in units of ppb months, and corrected for TIN and the date a person stopped drinking contaminated water (TSTOP). For cases, TAEVOC was calculated until TDIAG, if that date was earlier than TSTOP. An alternative exposure expression was TAEVCL. TAEVCL is equal to TAEVOC + the TAE for chloroform in city water. Finally, exposure was also expressed as DOSEVOC and DOSEVCL (equal to WATER times TAEVOC or TAEVCL), or as having lived in the contaminated area.

As indicated in Chapter 6, these options offer opportunities for exposure-effect analyses based on different viewpoints. However, DOSEVCL is the option most consistent with toxicology and risk assessment principles. In the analysis, exposure variables were used as 1) dichotomous variables, or 2) a variable with the values zero-low-high (the boundary between low and high was defined as the rounded median of all positive values). Exposure to VOCs via other routes was expressed as WASH (a measure of water use for bathing or showering, Equation 7 in Chapter 4), WASHVOC (WASH times the sum of VOC concentrations in the current water sample), and WASHVCL (WASH times the sum of the concentrations of VOCs + chloroform in city tap water).

F.4 Statistical Analysis

The start of the followup period (TFU1) was 1-1-1970, or TIN if a person moved to the area after that date. The close of the followup period (TFU2) was 1) the date of diagnosis for cases, and 2) the date of death or the date of interview for non-cases. The total followup time TFU was equal to TFU2 - TFU1.

The analysis included 1) a univariable analysis with fourfold tables, and 2) a multivariable analysis using Cox's proportional hazard model worked out by Harrell as a SAS-program PROC PHGLM (5). This model is intended for classic cohort studies where the exposure level is constant from TFU1 to TFU2. In the Battle Creek scenario, VOC levels changed over time, and people were not all exposed at the same time and for the same duration.

To solve this problem, PROC PHGLM was modified as described in Chapter 6. This revised version deals effectively with the fact that people enter the cohorts at different points in time, that most exposed people are initially unexposed, that exposure may stop before the end of followup, and that VOC exposure may occur simultaneously with exposure to chloroform from city water. An additional advantage is that it deals with real calendar time.

Common knowledge of the epidemiology of hypertension (3, 6, 7) indicated that the body mass index QI (Quetelet Index, which is weight in kg/height in m²), sex, blood lipids, and kidney functions were the most important risk factors to account for in the analysis. Thus, the multivariable models tested included one of the exposure variables, WATER, WASH (or WASHVOC - WASHVCL), age at TFU1, sex, QI, triglycerides (TRIG), and the first principal component KID1 of the triad blood urea nitrogen - uric acid - creatinine.

F.5 Results

There were 174 persons with some indication that hypertension might be present. Table F.1 shows a breakdown of this number by category of validity of the case diagnosis, and by eligibility for analysis. In total, 92 cases were eligible to enter the analysis; because some had missing values for variables used in the analysis, the number may be smaller for certain analytical models. None of the children younger than 16 years of age had blood pressures above the threshold set for abnormal blood pressure in children (3).

Table F.1 Potential cases of hypertension and cases accepted for analysis in exposed (Expos) and reference (Refer) cohorts.

Category of Diagnosis	Total Number HT>0		Ineligible cases TDIAG<1970		Ineligible cases TDIAG<TIN *		Cases eligible for analysis	
	Expos	Refer	Expos	Refer	Expos	Refer	Expos	Refer
HT=1	26	67	3	5	5	7	18	55
HT=2	7	15	1	0	1	1	5	14
HT=3	16	30	considered noncases					
HT=4	5	8	considered noncases					
Totals	54	120	4	5	6	8	23	69

* To prevent double counting, cases diagnosed before 1970 and before TIN were counted in the column TDIAG<TIN only.

Table F.2 depicts the frequency of reports of confirmed hypertension (HT=1-2) by source of information. The interview was the source with the highest yield of cases. On the other hand, it also yielded the highest number (N=40) of false positive reports, compared to N=17 if medical records were the only source. Blood pressure as an additional parameter significantly increased the number of positive findings. The number of false positives did not decrease significantly, however, as a normal pressure measurement did not rule out hypertension for people under treatment. Utilizing all four sources of information not only maximized the yield of cases, it also minimized the number of false negative findings.

Table F.2 Information sources for confirmed cases of hypertension.

Source	Number of Confirmed cases
Interview only	75
Medical records only	58
Clinical examination only	23
Clinical examination and/or interview	105
Clinical examination and/or medical records	94
Interview and/or medical records	91
BPMED use and/or interview	89
BPMED use and/or medical records	86
All four sources	115

Of 42 people with confirmed hypertension who had a positive interview *and* a positive medical record, seven were unable to recall the date of diagnosis at the interview, and three had an undated medical record. Of the 33 cases with dates from both the interview and the medical record, the date reported at interview was later than the medical record date in nine cases (range 2 - 98 months, average 42 months), three had no difference in these dates, and 21 persons reported dates earlier than those found in the records. Of the latter 21 cases, eight reported dates before 1970, a period not covered by the search for medical records. Ignoring these eight cases, 13 interviewees out of 33 (39%) reported a date of diagnosis earlier than the medical record date with a mean difference of 56 months (range 15 - 143 months), which may have been caused by a failure to retrieve earlier records. The above differences were unrelated to the exposure status.

For the population-based incidence rate (IR) of hypertension (categories HT=1 and HT=2) all cases diagnosed as of 1970 were counted (97 cases), regardless of whether the diagnosis was made before or after TIN. The IR is the number of cases divided by the total TFU (120,145 months) accrued by all people except those with hypertension before 1970. With a correction factor of 12 (to arrive at an annual IR), the IR for the period 1970-1985 for men and women combined is: $97 \times 12 \times 1000 / 120,145 = 9.69$ per 1000 population/year (95% Poisson confidence limits 7.86-11.82). The IR for the age group 15 and older (the youngest case was 16 years of age), is 11.41 (8.11-15.09). The prevalence rate as of July 1985 is the number of all valid cases (ignoring TDIAG) over the total number of people in the study, alive as of July 1, 1985, and is equal to 114/736 or 15.5 % (12.8-18.6). The prevalence for the population age 15 years and older is 20.1% (16.6-24.1). There is no significant differences in the sex-specific incidence and prevalence rates, as shown below in Table F.3.

Table F.3 Prevalence and incidence rates for hypertension.

Prevalence (all ages)		Incidence rate/1000/yr (all ages)	
Males	14.4% (10.6-19.0)	Males	9.33 (6.75-12.56)
Females	16.5% (12.7-21.0)	Females	10.00 (7.51-13.04)

Table F.4 summarizes the numbers of cases by exposure status: TAEVOC, TAEVCL, DOSEVOC and DOSEVCL and having lived in the contaminated area. The measures of association are the odds ratio (OR) and the relative risk (RR, based on the period of follow up) calculated as described in Chapter 6. Confirmed cases were eligible for analysis if they met the criteria for TDIAG; later than 1-1-1970 and later than TIN. The OR table for AREA is presented for the sole purpose of demonstrating the effect of a conventional approach in which the only eligibility criterion is that the uncorrected TDIAG is later than TIN (resulting in a larger number of cases).

Table F.4 ORs and RRs for hypertension (HYPER) and exposure expressed as a dichotomous (yes/no resident of contaminated area, yes/no positive value for TAE or DOSE variables) and a categorical variable (zero-low-high).

Total Population age 15 and over: Exposure expression is TAE or AREA								
Hypertens	Contam. AREA		Exposure = TAEVOC (VOCs)			Exposure = TAEVCL (VOC+Chl)		
	No	Yes	Exposed No	Exposed Yes	Exposed Low	Exposed No	Exposed Yes	Exposed Low
No	314	150	331	133	331	63	331	70
Yes	74	27	77	15	77	10	77	5
OR =	0.76		0.49		0.68		0.31	
p =	0.16		0.008		0.19		0.005	
RR =			0.54		0.79		0.33	
p =			0.01		0.30		0.004	

Total Population age 15 and over: Exposure expression is DOSE								
Hypertension	Contam. AREA		Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)		
	No	Yes	Exposed No	Exposed Yes	Exposed Low	Exposed No	Exposed Yes	Exposed Low
No	344	109	344	48	344	61	185	268
Yes	77	12	77	8	77	4	48	41
OR =			0.49		0.74		0.29	
p =			0.02		0.30		0.007	
RR =			0.56		0.88		0.33	
p =			0.03		0.45		0.03	

Table F.4 continued

Verona Exposed + Calhoun Reference 1977 - 1985: Exposure expression is TAE or AREA								
Hypertens	Contam. AREA		Exposure = TAEVOC (VOCs)		Exposure = TAEVCL (VOC+Chl)			
	No	Yes	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed High	
No	246	101	261	86	261	54	261	32
Yes	59	22	44	11	44	9	44	2
OR =	0.91		0.76		0.99		0.37	
p =	0.42		0.28		0.58		0.13	
RR =			0.82		1.04		0.43	
p =			0.35		0.52		0.17	

Verona Exposed + Calhoun Reference 1977-1985: Exposure expression is DOSE								
Hypertension	Exposure = DOSEVOC (VOCs)		Exposure = DOSEVCL (VOC+Chl)					
	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed High			
No	271	66	271	39	271	27	114	223
Yes	44	9	44	6	44	3	22	31
OR =		0.84		0.95		0.68		0.72
p =		0.41		0.56		0.39		0.17
RR =		0.94		1.02		0.81		0.79
p =		0.52		0.54		0.50		0.25

- 1) p-values were estimated by a 1-tail Fisher's exact test.
- 2) The numbers of people and cases in the DOSE tables are somewhat smaller than the total number in the TAEVOC tables, since people with missing data for WATER (and thus for DOSE) were deleted.
- 3) The threshold between Low and High exposure level is determined by the median of all non-zero values. "Low" and "High" are relative terms for the analysis only. They do not have a toxicological connotation!

Table F.4 shows a statistically significant excess of disease in the reference cohort, as reflected in ORs and RRs below unity for all exposure expressions. Limiting the analysis to the Verona exposed and the Calhoun reference cohort, for reason of better exposure data, did not change this outcome.

The results of the multivariable analysis for TAEVCL and DOSEVCL are shown in Table F.5. The exposure variables have been transformed into their natural logarithms (LTAEVCL and LDOSEVCL). In the model with DOSEVCL, WATER is deleted as it is already incorporated in the DOSE variable, while WASH is replaced by WASHVCL. In these models, exposure was calculated as TAE and DOSE accumulated since TFU1 up to the end of the index period, as described in Chapter 6.

All coefficients of the exposure variable were negative, that is, the dose-response effect is negative and $OR < 1$. The exposure variables, WATER, and WASH did not contribute in any degree to the model fit (partial $R=0$). By far the largest contributors to the R-statistic were age at TFU1 and the body mass index QI. The results of the analysis using TAEVOC and DOSEVOC were similar, and were, therefore, not shown in the table.

Table F.5 Results of proportional hazard analysis, blocking for period. The values shown are the coefficient (beta) and p-value (p) of the variables in the model. AGE1 is the age at the start of the followup; N is the number of cases with no missing values for either of the variables; R is the R-statistic akin to the multiple correlation coefficient (4).

Cohorts involved	Model 1 (N=81, R=0.319)			Model 2 (N=82, R=0.319)		
	Variable	beta	p	Variable	beta	p
All cohorts	LTAEVCL	- 0.1093	0.0241	LDOSEVCL	- 0.058	0.1279
	WATER	- 0.0063	0.7802			
	WASH	0.0116	0.2545	WASHVCL	- 0.0000	0.4817
	AGE1	0.0554	< 0.0001	AGE1	0.0511	< 0.0001
	SEX	- 0.5176	0.0580	SEX	- 0.4977	0.0555
	KIDN1	0.3059	0.0027	KID1	0.2811	0.0026
	QI	0.0694	< 0.0001	QI	0.0656	< 0.0001
	TRIGLYC	0.0009	0.0258	TRIGLYC	- 0.0009	0.0278
Verona exposed and Calhoun reference	Model 3 (N=46, R=0.325)			Model 4 (N=47, R=0.321)		
	LTAEVCL	- 0.0807	0.2380	LDOSEVCL	- 0.0301	0.5584
	WATER	0.0061	0.8421			
	WASH	0.0251	0.0450	WASHVCL	- 0.0000	0.8338
	AGE1	0.0533	< 0.0000	AGE1	0.0474	< 0.0000
	SEX	- 0.6630	0.0679	SEX	- 0.5978	0.0785
	KID1	0.4110	0.0039	KID1	0.3358	0.0056
	QI	0.0802	< 0.0001	QI	0.0730	< 0.0001
TRIGLYC	- 0.0008	0.1313	TRIGLYC	- 0.0008	0.1123	

Multivariable analysis with other TAE expressions as described in Chapter 6 did not change these results. Since several covariates reflect a status at TFU2 rather than TFU1, the analyses were repeated with age as the only covariate. Again, no changes were seen in the negative direction of the exposure variables. In addition, correlation tests for SP or DP (systolic or diastolic blood pressure) with TAE and DOSE yielded consistently negative correlation coefficients. Multiple linear regression with these exposure expressions as the dependent variable, and SP or DP as the regressor with the risk factors in Table F.5 plus a number of

other potential predictors (see next paragraph) as covariates, showed a consistently negative coefficient for the exposure variable.

For a better evaluation of the risk factors of hypertension other than exposure to VOCs, the data were analyzed with a conventional PROC PHGLM (no partitioning of the followup period) with TFU1 starting at age 21 (to control for age, and to restrict the analysis to adults), and ignoring the exposure variables, WASH, and WATER. In addition to the covariates listed in Table F.5, the risk factors tested were serum levels of the pesticide DDT, PCB and PBB (polychlorinated and polybrominated biphenyls), cholesterol, income and education, occupational exposure to VOCs, pesticides, or other chemicals, and liver enzymes. The full equation for significant risk factors, while controlling for age (TFU1 was set for all people at age 15), is shown below. In this equation, CASE is any valid case of hypertension (no eligibility criteria for TDIAG), SCHOOL is a measure of the education level (range 1-6), WORK1 represents occupational exposure to VOCs, URIC is serum uric acid, and LDDT is the logarithm of DDT. The p-values associated with the coefficients ranged from 0.02 to < 0.0001.

MODEL : (R = 0.22, -2log L = 810 and model p < 0.001)

TFU = 0.062 QI + 0.319 SCHOOL + 0.544 WORK1 + 0.155 URIC + 0.001 TRIG - 0.274 LDDT

F.6 Discussion

The main question underlying the Battle Creek Health Study is whether exposure to VOCs in drinking water caused any observable health effect. With respect to hypertension, this question can be answered conclusively. Exposure to VOCs at the prevailing levels can be ruled out as a causal factor in the development of hypertension. Also, such exposure has not had an increasing effect on either the SP or DP. On the contrary, blood pressure decreased with increasing TAE or DOSE. The absence of an adverse effect of exposure did not result from too small a study size; this is shown by the consistently negative direction of the association of exposure with the risk of hypertension -- across analytical models, strata, case definitions (HT=1, HT=1-2, HT=1 through 4), and exposure expressions. Finally, although WORK1 was a statistically significant risk factor if tested in a model without exposure variables, it proved insignificant when added to the models shown in Table F.5 ($p > 0.3$).

In summary, all findings clearly point to a statistically significant excess of people with hypertension in the reference cohort. There are no biological data supporting a "protective" effect of VOCs. If exposure to VOCs was disregarded in the analysis, the risk factors for hypertension were increasing age, obesity (QI), elevated triglycerides, and impaired kidney function (in particular expressed as elevated serum values of uric acid), which is in full agreement with known epidemiologic facts. Of these risk factors, uric acid had the least impact on the risk of disease. This may perhaps be explained by the use of BPMED. Gender was a necessary and significant covariate in the models tested; although the IR and prevalence were the same for men and women, blood chemistry and occupational exposure were different. Another statistically significant risk factor was family income, but only if tested in the absence of SCHOOL. The association may be indirectly causal (through health habits), or circumstantial, because persons with higher income and education levels are more likely to have medical checkups.

Alcohol abuse has been found to increase the arterial pressure (3, 8, 9), but the association with hypertension remains to be established (9). In this study, alcohol consumption was not associated with hypertension or arterial pressure. The lack of such association may, however, also indicate that the quantitative information on alcohol consumption, obtained during interview, was

inaccurate, or that this effect of alcohol is associated with a minimum exposure threshold. Due to time constraints, this possibility has not been investigated.

Although most diabetics have hypertension, the association between these disorders is probably indirect through a wide variety of environmental factors (10). Whether or not hypertension follows diabetes (or vice versa) remains an issue to be investigated (3, 7, 11). It has been established that certain types of blood pressure controlling drugs may affect the glucose metabolism (3, 7), and Ritz and coworkers (7) speculated that this might deteriorate into diabetes. On the other hand, they also indicated that renal lesions in diabetes might invoke a hypertensive reaction. In this study, no statistical association was found between these disorders, but this may have resulted from the small number of diabetics, or from the mixed character of the population of cases. Diastolic hypertension and isolated systolic hypertension have different risk factors, and the epidemiology of borderline hypertension differs from what is known of established hypertension (3, 9, 12). The number of hypertensives in this study was too small, however, for an analysis by type of hypertension. It is also beyond the scope of the current study to expand the analysis of risk factors of hypertension beyond what has been reported above.

PCB was not found associated with either hypertension or SP and DP. A positive association with blood pressure, in particular DP, was reported by Kreiss *et al* (13), who did not report on the relation with hypertension *per se*. Kreiss did not include QI in the model, which in the current study was the most important predictive variable. Kreiss *et al* did not describe how the population sample was obtained, and their definition of hypertension included people with abnormal blood pressures only. It is unclear whether people successfully (becoming normotensive) treated with drugs were included or not. Since no description was given of the analytical model used, it is impossible to evaluate the appropriateness of the analysis.

It should be recognized that any analysis of historical and current data has the implicit potential of yielding erroneous results, since current laboratory values may not be directly related to past levels. The results shown in Tables F.4 and F.5 are, thus, presented with the assumption such an relation does exist. The covariates in Table F.4 were intentionally selected for the likelihood that they were indicative of the past health status. Occupation and education level in adults vary little over time. In this study, about two thirds of the cases underwent the clinical examination less than 10 years since their hypertension was diagnosed. Within this short period, it is reasonable to assume that the current levels of QI, kidney function, and triglycerides are well correlated with past values. Following the same line of reasoning, more volatile parameters, such as DP, SP, and liver enzymes, may not be suitable variables for historical analysis.

In analyzing associations of hypertension and VOCs, it is important to have accurate data not only on exposure, but also on health status. Although quality assurance and control procedures were applied to the limit of what was practically feasible (raising the budget by about one third), assessment of the historical health status cannot be more accurate than are the stored data. From the viewpoint of risk assessment, the date when a disease developed as a result of exposure should be when the development reached an irreparable stage. In this study, a lag time of 1 year in TDIAG was taken into account. However, the real lag time remains unknown. Knowledge of the genesis of hypertension, in particular following chemical exposure, is minimal. Hence, the use of a lag time of 1 year is merely an acknowledgement of the problem rather than a best estimate. Evaluation of disparities between the reported (interview) and recorded (medical file) dates of diagnosis revealed no association of errors with the exposure status. Hence, given the outcome of the analysis, and the direction of the coefficients of the exposure variables, the use of a lag time longer than 1 year is unlikely to change the outcome of the study, and it would have had the disadvantage of reducing the number of eligible cases.

Another aspect of accuracy of data is the completeness of the case-finding survey. In the literature review, no example was found for an IR measured in a total (all ages and both sexes, current and past residents) community followed up for a sufficiently long period. The observed annual IR of 9.7/1000 population of all ages and both sexes (the difference in sex-specific rates was small and statistically insignificant) is, therefore, new information. The prevalence measured in the National Health and Nutrition Examination Survey 1976-1980 (NHANES) for the age-group 18-74 was 17.4% for white males and 18% for females (14). The prevalence rate found in this study (recomputed for this age-group) is 20.7% and 21.0%, respectively. These higher rates (the difference is statistically insignificant) may be attributable to the fact that NHANES covered the non-institutionalized population only, while the current study included the entire population. Further, the NHANES did not screen medical records and did not inspect the medication taken for blood pressure lowering drugs. In any event, the higher rates in the current study render it very unlikely that incompleteness of the data base might have affected the outcome of the statistical analysis.

From a public health viewpoint, the major important finding in this study is that a relatively large proportion of hypertensive people were 1) unaware that they had the disease, 2) without medication, or 3) apparently over-medicated. This is reflected in the following facts:

- 1) Out of 115 people with confirmed hypertension (HT=1-2), 40 were unaware of it and 26 of these had no BPMED. There were 23 cases who did not know they had hypertension and had no BPMED and no positive medical record; nine of these had an established hypertension (two as high as 143-116 and 215-101 mm).
- 2) Out of 51 people with confirmed hypertension and a positive medical record and BPMED use, 13 responded negatively when asked about the presence of the disease. Of these 13 people, two had pressures above the threshold of established hypertension (167-90 and 176-118 mm) while using BPMED.
- 3) Out of 92 people on BPMED (among whom were all category HT=4 people), 41 had no medical record of hypertension, and 14 of these said they had no hypertension. Some of these 14 people might have had BPMED for reasons other than hypertension. Among these 14, were four adults who had a clear hypotension (arbitrary cutoff point of SP < 100 mm, mean pressure 92-64 mm). This is suggestive of hypotension from over medication, which may reflect poor medical followup. A much larger number of hypotensives would result if the cutoff point was set at a higher level to account for an increasing threshold with increasing age.

Thirteen out of 51 people (25%) who were on BPMED and had a positive medical record, were unaware of having hypertension. This may be due to recall deficiency, and also to less than optimal communication between physician and patient. This is supported by the finding that 14 people, out of 92 using BPMED, had no medical record of hypertension and were unaware that they used drugs with a blood pressure lowering effect, even though 30% of them were clearly hypotensive, apparently as a result of their medication. Out of 115 people with confirmed hypertension, three were unaware (though two had a medical record) that they had a serious and untreated hypertension. That the public's awareness of hypertension is improving, though still insufficient, is clear from the data compiled by the National Heart, Lung, and Blood Institute (NHLBI), and reviewed by Lenfant (15). Table F.6 shows data on awareness, medication, and effect of medication from the NHBLI and from the current study.

Table F.6 is reflective of an awareness in the study population much lower than the national average, despite the higher percentage of BPMED users. The higher medication rate explains the higher "control" rate, but both are misleadingly high. As discussed above, a substantial number of people may have been using BPMED while not hypertensive (e.g., all of category HT=4). There

is considerable controversy about whether borderline hypertension (particularly in those under 50 years of age) should be treated with pharmacotherapy (3). Increased blood lipids are among the known adverse effects (3), and an excess of high values of blood lipids and an excess of hypertensives were found in the reference cohort. This might be coincidental, and there are differences between the data base of the NHBLI and that of the current study. Thus, Table F.6 should be viewed with caution. Information on BPMED use in the NHBLI population was obtained by interview, not by inspection and verification of the drugs. Further exploration of this issue is justified by: a high medication rate in the study population (compared to the national average), in conjunction with an observed excess of abnormal laboratory values regardless of the exposure status (see Table 8.2, Chapter 8); and the risk that over-medication or unnecessary medication may lead to adverse effects.

Table F.6 Public awareness and medical control of hypertension (defined as SP > = 160 mm or DP > = 95 mm, or on medication).

Source of data	Percent awareness *	Percent on medication	Percent control **
NHBLI study 1982-84	85	74	57
Current study 1985	67	88	83

* NHBLI: percent of hypertensives informed by physician. Battle Creek: percent of those who said they had hypertension.

** Percent of cases on medication resulting in SP < 160 mm and DP < 95 mm. The denominator may have included some normotensives, as the NHBLI definition accepted BPMED use as "evidence" of disease, and BPMED might have been prescribed for reasons unrelated to hypertension.

The degree of inability of respondents to accurately recall when their disease was first diagnosed is more or less comparable to what has been observed in the study population with respect to diabetes (Appendix E). There was a failure to recall any date by 11 out of 75 (15%) cases in category HT=1-2 who knew they had hypertension. Nine out of 42 (37%) people with confirmed hypertension and a positive medical record and a positive response mentioned a date of diagnosis up to 8 years later than found in their records. This shows the need in health studies for quality control procedures which go beyond controlling laboratory performance. Since the majority of people who mentioned erroneous dates of diagnosis were age 30 to 65 years, mental deficiency is an unlikely explanation for the inaccuracy in the recall. No recall differences were observed between residents of exposed and comparison areas, despite publicity about the toxicity of VOCs.

F.7 Conclusion and Summary

Following the detection of groundwater contaminated with chlorinated volatile organic compounds (VOCs), a retrospective cohort study was conducted involving 749 people with an average followup period of 13 years, and a ratio of exposed to unexposed subjects of approximately 1:2. The design of the study followed the principles of risk assessment for the assessment of the health risk from toxic chemicals. Although the study was not designed to study hypertension specifically, the presence of 174 possible cases of hypertension warranted a detailed analysis of this disease. After validation of the available health information, 115

confirmed cases of established and borderline hypertension remained, equal to a prevalence as of July 1985 of 15.5% and an incidence rate of 9.7/1,000/yr for the total population (all ages and both sexes). There was no significant difference in the sex-specific rates. The results of the study warrant the following conclusions:

- 1) Exposure to VOCs in drinking water and occupational exposure to VOCs are not related to the risk of developing hypertension.
- 2) The risk factors for hypertension include age, obesity, an impaired renal function, higher education level, and diabetes.
- 3) Alcohol consumption, occupational exposure to various chemicals, DDT, PCB, and PBB (at the prevailing serum levels), smoking, and liver function tests have not shown an association with the occurrence of hypertension, but the limited time available did not allow a more detailed analysis, taking into account threshold levels, concurrent diseases, and different types of hypertension.

Other important findings from the study are: a) an historical survey for medical data limited to an interview and clinical examination or with a medical record survey as a second source of information is bound to underestimate the disease incidence and prevalence, and does not permit adequate control of the quality of the health data compiled; b) there is a substantial inaccuracy in the dates of diagnosis of hypertension reported by the interviewees; c) one third of the study population with established hypertension were unaware of their condition, higher than the percentage found in a national population sample, and there is evidence that patient-physician communication and the utilization of medical services can be improved.

F.8 References

1. Wright, B. M. and Dore, C. F. A random-zero sphygmomanometer. *Lancet* 1: 337-338 (1970).
2. Maxwell, M. H., Schroth, P. C., Waks, A. U., et al. Error in blood pressure measurement due to incorrect cuff size in obese patients. *Lancet* 2: 33-36 (1982).
3. The 1984 Report of the Joint National Committee on detection, evaluation, and treatment of high blood pressure. *Arch. Intern. Med.* 144: 1045-1057 (1984).
4. Freni, S. C. and Phillips, D.L. Estimation of the time component in the movement of chemicals in contaminated groundwater. *Environ. Health Perspect.* 74: 211-221 (1987).
5. Harrell, F. E., Jr. The PHGLM procedure. In: *SUGI Supplemental Users Guide*, SAS Institute Inc., Cary NC, 5th edition pp. 438-466 (1986).
6. Schmieder, R. E., Messerli, F. H., Ruddel, H. Risks for arterial hypertension. *Cardiol. Clin.* 4: 57-66 (1986).
7. Ritz, W., Hasslacher, C., Tschöpe, W., et al. Hypertension in diabetes. *Contr. Nephrol.* 54: 77-85 (1987).
8. Klatsky, A.L. The relationship of alcohol and the cardiovascular system. *Ann. Rev. Nutr.* 2: 51-71 (1982).
9. World Health Organization. *Primary Prevention of Essential Hypertension. Technical Report Series Nr. 686*, WHO, Geneva, 1983.
10. World Health Organization Study Group. *Diabetes Mellitus. Technical Report Series 727*, World Health Organization, Geneva, 1985.
11. Herman, W. H., Teutsch, S. M., and Geiss, L. S. (prepared for The Carter Center). Closing the gap: the problem of diabetes mellitus in the United States. *Diabetes Care* 8: 391-406 (1985).
12. Lichtenstein, M. Isolated systolic hypertension. *South. Med. J.* 78: 972-978 (1985).
13. Kreiss, K., Zack, M. M., Kimbrough, R. D., et al. Association of blood pressure and polychlorinated biphenyl levels. *JAMA* 245: 2505-2509 (1981).
14. National Center for Health Statistics. *Blood Pressure Levels in Persons 18-74 Years of Age in 1976-1980, and Trends in Blood Pressure from 1960 - 1980 in the United States. Data from the National Health Survey Series 11, No. 234*. U.S. Department of Health and Human Services, DHHS Publication no. (PHS) 86-1684, Hyattsville, MD, July 1986.
15. Lenfant, C. Advancements in meeting the 1990 hypertension objectives. *MMWR* 36: 144, and 149-151. (1987)

Appendix G

**Gall Bladder Disease and Exposure to Chlorinated
Volatile Organic Chemicals in Drinking Water**

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

G.1 Introduction

In 1981, the aquifer at the Verona Well Field, Battle Creek, Calhoun County, Michigan, was found to be contaminated with chlorinated, short-chain volatile organic chemicals (VOCs). Most of the city wells, and a large number of residential wells were affected. Subsequently, VOCs were also detected in the water of a number of private wells in Springfield, which is adjacent to Battle Creek, and in Dowagiac, Cass County, Michigan. In 1984, Phase I of a two-phase epidemiologic study was initiated, basically to collect demographic data to form exposed and reference study populations for Phase II, a retrospective followup study, which started in July 1985. In Phase II, information was collected on morbidity and exposure for exposed and reference cohorts. The study involved 251 residents from contaminated areas and 498 from reference neighborhoods. The groups were matched for sex, age, family size, and neighborhood. Since 94% of the study population were white, race was not considered in the analysis.

Since the study focused on the potential effects of a predefined exposure, the principles of risk assessment for chemicals were applied. These included:

- 1) Exposure must be quantified as a dose and in time.
- 2) A dose-response effect must be determined.
- 3) A disease is attributable to exposure only if the exposure preceded the disease.
- 4) Quality control procedures must be extended to the quality of the health data.
- 5) Uncertainties in the data and methods, and their effect on the study results, must be properly addressed.

Initially, there was no toxicologic or epidemiologic indication that gall bladder disease was associated with exposure to VOCs. Preliminary analysis of the crude study results, however, showed a substantial number of possible cases of gall bladder disease. Since little is known about the causation of this disease, and a role for VOCs could not be precluded, further analysis was considered worthwhile and feasible.

G.2 Health Assessment and Case Definition

All hospitals serving the study area (Appendix B) were canvassed in a survey for the medical records of all study participants. In addition, physicians and hospitals specifically mentioned by the study participants were contacted to retrieve medical records. Copies of all death certificates were obtained. All medical records and death certificates were screened for information on diseases occurring as of 1970. In-person interviews and clinical examinations were conducted. Blood and urine samples were collected from (fasting) participants. There was no provision for X-ray examination.

The questionnaire was designed to maximize the recall of the respondent by repeating some questions in different formats (Appendix A). The questions relevant to gall bladder disease asked about: any current disease or illness, a history of gallstones, hospitalization for gallstones prior to and since 1970, and (at the close of the interview) about any disease or condition not earlier mentioned. Each question also asked for the date of diagnosis and name and address of hospital and physicians. Names of drugs and the dates they were prescribed were recorded. This multisource approach to collecting data on the individual health status was chosen to detect and correct recall failure, erroneous responses, and poor records. In the case of gall bladder disease, only interview data and records were useful.

Without prior knowledge of their exposure status, people with any indication of possible gall bladder disease were categorized according to the likelihood that gall bladder disease was truly present. Cholecystitis and cholelithiasis were combined, as these disease entities usually co-exist and there was no information allowing separation. (Actually, there were only five cases in which the medical record stated chronic inflammation and not stones. In all these cases, the interviewee mentioned gallstones). The disease categories are listed below:

GALL=1:

A positive medical record, regardless of the interview response.

GALL=2:

A positive interview; hospitalization for the disease was mentioned, although no medical record was found. Rationale: if the diagnosis was made prior to 1970, and if recurrence in later years did not require medical attention, no medical record would have been found, as the search for records did not cover those years.

GALL=3:

A positive interview, but the respondent indicated no hospitalization for the disease, and the medical record was negative.

A diagnosis was considered valid if GALL = 1. Since gall bladder disease cannot be validated on the basis of the laboratory tests, and failure to retrieve a medical record could not be precluded, an alternative case definition included GALL=2 as well. People with GALL=3 were considered noncases, since no medical record *and* not having been hospitalized *and* not having used chenodiol indicated that cholelithiasis was unlikely. Cases were eligible to enter the analysis if the date of diagnosis (TDIAG) was later than the date of moving to the study area (TIN) *and* later than 1-1-1970. In the analysis of the association of exposure with GALL=1, GALL=2 cases were deleted, as it was uncertain that these people were indeed noncases. TDIAG was corrected for a lag time arbitrarily set at one year, that is, moved back one year as described in Chapter 6.

G.3 Exposure Assessment

The individual exposure level, accumulated during the period of residency in the study area, was estimated as TAE (total accumulated exposure), equal to the area under the $\log C = T$ curve described in Appendix C and elsewhere (1). In this notation, $\log C$ is the natural log of the concentration C of a given chemical in the well water sample (in parts per billion or ppb), and T is the date of sampling expressed in months since 1-1-1970. TAEVOC is the sum of chemical specific TAEs in units of ppb-months, corrected for TIN and the date a person stopped drinking the contaminated water (TSTOP). For cases, TAEVOC was computed until TDIAG if this was earlier than TSTOP. An alternative exposure expression was TAEVCL, equal to TAEVOC + the TAE for chloroform in city water. Other exposure expressions were 1) DOSEVOC and DOSEVCL, equal to WATER times TAEVOC and TAEVCL, respectively, and 2) having lived in the contaminated AREA. These options for defining exposure offer opportunities for exposure-effect analyses from different viewpoints. However, DOSEVCL is the option most consistent with toxicology and risk assessment principles. In the analysis, exposure variables were used as dichotomous variables, variables with the values zero-low-high (the boundary between low and high is determined by the rounded median of all positive values), or as continuous variables.

G.4 Statistical Analysis

The start of the followup period (TFU1) was 1-1-1970, or TIN if that date was later. The close of the followup period (TFU2) was TDIAG for cases, and the date of death or the date of interview for non-cases. The total followup time $TFU = TFU2 - TFU1$. Since no gall bladder disease occurred in people under age 20 (the youngest was age 22 years), the analysis was done for people aged 21 and over. The analysis was performed as: 1) a univariable analysis involving fourfold tables, and 2) a multivariable analysis using Cox's proportional hazard model, worked out by Harrell as a SAS program PROC PHGLM (2). This model is intended for classic cohort studies, that is, exposure is constant from TFU1 to TFU2. In this study, the VOC level was time-dependent, and people were not all exposed at the same time or for the same duration.

The modified PROC PHGLM, described in Chapter 6, effectively deals with this and other problems: people entered the study at different points in time; most people were initially unexposed; exposure may cease before TFU2; and exposure to VOCs may be followed by exposure to chloroform from city water. Another advantage is that the model deals with calendar time. The basic model tested included, as predictive variables, SEX, plus one of the expressions of exposure mentioned earlier -- WATER (in models with TAEVOC or TAEVCL) or WASH (a measure of water use for bathing and showering, Equation 7, Chapter 4). The small number of eligible cases prohibited the use of more covariates.

G.5 Results

The collected health information yielded 55 potential cases of gall bladder disease. Table G.1 categorizes these reports according to 1) the likelihood that the report refers to a true case, and 2) the reason for rejecting a case for the analysis of exposure-disease associations.

Table G.1 Potential cases of gall bladder disease and cases for analysis in exposed (Expos) and reference (Refer) cohorts.

Category of Diagnosis	Total number GALL>0		Ineligible cases TDIAG<1970		Ineligible cases TDIAG<TIN*		Cases eligible for analysis	
	Expos	Refer	Expos	Refer	Expos	Refer	Expos	Refer
GALL=1	15	19	1	0	4	6	10	13
GALL=2	4	10	1	3	1	4	2	3
GALL=3	0	7	considered noncases					
Totals	19	36	2	3	5	10	12	16

* To prevent double counting, cases diagnosed before 1970 and before TIN were counted in the column TDIAG<TIN only.

For the population-based incidence rate (IR) and prevalence of gall bladder disease, all cases of GALL=1 diagnosed since 1-1-1970 were counted, regardless whether the diagnosis was made before or after TIN. The IR is the number of cases divided by the total TFU (124,349 person-months) accrued by all people except those who had gall bladder disease before 1970. With a correction factor of 12 (conversion of months to years), the IR for the period 1970-1985 for men and women combined is $31 \times 12 \times 1000 / 124,349 = 2.99$ per 1000 population/yr (95% Poisson confidence limits of 2.03-4.25). Expanding the case definition to GALL=1 or 2 yields an

IR = $40 \times 12 \times 1000 / 123507 = 3.89/1000/\text{yr}$ (2.78 - 5.29). The prevalence as of July 1985 is the number of cases over the number of people in the study *alive* by that date, and is equal to 33/736 or 4.5% for GALL=1 (Poisson 95% confidence limits 3.0-6.3). Expanding the case definition to GALL=1 or 2 raises the prevalence to 46/736 or 6.3% (4.6-8.3). The gender-specific prevalence and IR are:

Table G.2 Gender-specific prevalence and incidence Rate for Gall Bladder Disease.

Prevalence (all ages):			Incidence rate/1000 (all ages)/yr		
males	GALL=1:	2.6% (1.2-5.0)	males	GALL=1:	1.66 (0.72-3.27)
females	GALL=1:	6.1% (3.9-9.0)	females	GALL=1:	4.15 (2.63-6.23)
males	GALL=1&2:	3.2% (1.6-5.8)	males	GALL=1&2:	1.87 (0.85-3.54)
females	GALL=1&2:	8.9% (6.2-12.3)	females	GALL=1&2:	5.67 (3.85-8.04)

Tables G.3 and G.4 depict the distribution of cases and noncases by exposure status for both case definitions and for the total population age 21 and over. The measures of association are the odds ratio (OR) and the relative risk (RR), calculated for eligible valid cases as described in Chapter 6, with p-values from a Fisher's one-sided exact test. The tables show a deficit of gall bladder disease in the exposed cohort (OR and RR below unity) with a consistently negative dose-response effect, for both case definitions and the exposure expressions based on TAE and DOSE. The deficit of cases in the exposed cohort is statistically significant for TAEVCL and DOSEVCL. Restricting the analysis (for reasons of better exposure data) to the Verona exposed and the Calhoun county reference cohort changed little in the findings.

If exposure is defined as having lived in the contaminated area, the OR is above unity. The number of cases in the table for this exposure expression is larger than in the other OR tables, because no eligibility criterion applied other than that the *uncorrected* TDIAG was later than TIN. This approach, which follows the convention of nearly all environmental studies, ignores the role of an exposure assessment in concordance with risk assessment principles. That is, no assessment is made of when exposure began relative to TDIAG, and whether individual exposure actually took place. The OR table for this exposure is presented solely for the purpose of comparing the result of a conventional approach with more meaningful tables.

In the multivariable analysis of exposure-effect associations, the modified PROC PHGLM was used with the natural logarithm of TAE and DOSE. In the model with DOSE, WATER was deleted as it was already incorporated in DOSE, and WASH was altered into the variable WASHVCL, equal to WASH x CS (CS = sum of concentrations of VOCs in the current water sample) or WASH x 6 (6 = the chloroform concentration for residents on city water). Table G.5 shows the results of this analysis for case definition GALL=1 or 2, and for TAEVCL and DOSEVCL.

Table G.3 OR and RR for gall bladder disease (GALL) in people over age 20. Exposure is expressed as a dichotomous (yes/no resident of the contaminated area; yes/no positive value for TAEVOC or TAEVCL) or a categorical variable (no, low, or high TAEVOC or DOSE). Cases of GALL=1 are confirmed by a positive medical record.

Total Population 1970-1985: Exposure expression is AREA or TAE. Case definition GALL = 1.

Gall	Contam. AREA		Exposure = TAEVOC (VOCs)				Exposure = TAEVCL (VOC+Chl)							
	No	Yes	Exposed No	Exposed Yes	Exposed No Low	Exposed No High	Exposed No Yes	Exposed No Low	Exposed No High					
No	310	146	317	125	317	59	317	66	151	291	151	223	151	68
Yes	15	13	17	6	17	3	17	3	13	10	13	7	13	3
OR =	1.84		0.90		0.95		0.85		0.40		0.37		0.51	
p =	0.09		0.52		0.62		0.54		0.02		0.03		0.23	
RR =			0.92		1.05		0.82		0.46		0.43		0.55	
p =			0.54		0.56		0.52		0.05		0.05		0.26	

Total Population 1970-1985: Exposure expression is DOSE.

Gall	Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)								
	Exposed No	Exposed Yes	Exposed No Low	Exposed No High	Exposed No Yes	Exposed No Low	Exposed No High					
No	324	107	324	53	324	54	170	261	170	204	170	57
Yes	16	5	16	3	16	2	12	9	12	7	12	2
OR =		0.95		1.15		0.75		0.49		0.49		0.50
p =		0.58		0.52		0.52		0.08		0.10		0.29
RR =		0.99		1.22		0.77		0.55		0.56		0.54
p =		0.61		0.47		0.53		0.13		0.16		0.32

Verona Exposed and Calhoun County Reference Cohorts 1977-1985

	TAEVOC			TAEVCL			DOSEVOC			DOSEVCL		
	0-1	Low	High	0-1	Low	High	0-1	Low	High	0-1	Low	High
OR =	1.20	1.52	0.73	0.36	0.36	0.37	0.98	1.64	0	0.36	0.41	0
p =	0.49	0.37	0.61	0.06	0.07	0.31	0.62	0.34	0.32	0.05	0.09	0.13
RR =	1.24	1.57	0.76	0.39	0.39	0.40	1.04	1.67	0	0.48	0.56	0
p =	0.46	0.35	0.63	0.07	0.07	0.34	0.58	0.31	0.35	0.14	0.21	0.22

Table G.4 OR and RR for gall bladder disease (GALL) in people over age 20. Exposure is expressed as a dichotomous (yes/no resident of the contaminated area; yes/no positive value for TAEVOC or TAEVCL) or a categorical variable (no, low, or high TAEVOC or DOSE). Cases of GALL=1&2 had a positive medical record, or indicated at interview that they had been hospitalized for the disorder (no medical record was found).

Total Population 1970-1985: Exposure expression is AREA or TAE. Case definition GALL = 1.

Gall	Contam. AREA		Exposure = TAEVOC (VOCs)				Exposure = TAEVCL (VOC+Chl)							
	No	Yes	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed No	Exposed High	Exposed No	Exposed Low	Exposed No	Exposed High		
No	294	142	317	125	317	59	317	66	151	291	151	223	151	68
Yes	27	15	21	7	21	3	22	4	17	11	17	7	17	4
OR =	1.15		0.85		0.77		0.92		0.34		0.28		0.52	
p =	0.40		0.45		0.47		0.57		0.003		0.003		0.19	
RR =			0.89		0.87		0.91		0.41		0.35		0.59	
p =			0.49		0.56		0.56		0.02		0.01		0.24	

Total Population 1970-1985: Exposure expression is DOSE.

Gall	Exposure = DOSEVOC (VOCs)			Exposure = DOSEVCL (VOC+Chl)		
	Exposed No	Exposed Yes	Exposed No	Exposed Low	Exposed No	Exposed High
No	324	107	324	53	324	54
Yes	20	6	20	3	20	3
OR =		0.91		0.92		0.90
p =		0.53		0.59		0.58
RR =		0.97		1.00		0.95
p =		0.59		0.61		0.61

Verona Exposed and Calhoun County Reference Cohorts 1977-1985

	TAEVOC			TAEVCL			DOSEVOC			DOSEVCL		
	0-1	Low	High	0-1	Low	High	0-1	Low	High	0-1	Low	High
OR =	1.09	1.39	0.66	0.32	0.32	0.32	0.90	1.50	0	0.32	0.37	0
p =	0.55	0.42	0.57	0.03	0.03	0.25	0.58	0.38	0.29	0.03	0.05	0.10
RR =	1.14	1.45	0.70	0.35	0.37	0.35	0.97	1.55	0	0.35	0.40	0
p =	0.51	0.39	0.59	0.04	0.05	0.29	0.63	0.35	0.33	0.04	0.06	0.13

Table G.5 Results of proportional hazard analysis, blocking for period. The values shown refer to the coefficient (beta) and the p-values (p) of the variables in the model. R is the R-statistic akin to the multiple correlation coefficient in the normal setting (2).

Cohorts involved	Model 1 (N=22, R=0.09)			Model 2 (N=22, R=0.165)		
	Variable	beta	p	Variable	beta	p
All cohorts Case=1 if GALL=1&2	LTAEVCL	- 0.0764	0.2392	LDOSEVCL	- 0.0902	0.0986
	WATER	0.0171	0.6142	WASHVCL	0.0000	0.0161
	WASH	0.0059	0.5444	SEX	- 1.4428	0.0096
	SEX	- 1.3801	0.0138			
	Model 3 (N=13, R=0.0)			Model 4 (N=13, R=0.136)		
Verona exposed and Calhoun reference Case=1 if GALL=1&2	LTAEVCL	- 0.1243	0.1853	LDOSEVCL	- 0.1313	0.0977
	WATER	- 0.0102	0.8668	WASHVCL	- 0.0000	0.7319
	WASH	0.0077	0.4794	SEX	1.5646	0.0420
	SEX	- 1.5122	0.0516			

As shown in Table G.5, all coefficients of TAE or DOSE were negative (for DOSE statistically significant), which concurs with the negative ORs and RRs and the negative dose-response effect depicted in Tables G.3 and G.4. The variables WASH (or WASHVCL) and WATER showed inconsistent associations with the occurrence of disease, and always with large p-values. All multivariable analyses were also repeated for the more restricted case definition GALL=1, and for models with TAEVOC and DOSEVOC (log transformed). The results did not change to a significant degree from those shown in Table G.5. In still other models, SEX was replaced by QI (the body mass or Quetelet Index) or age at TFU1, again without significantly different results. In summary, neither the univariable nor the multivariable analyses showed any evidence of a risk of gall bladder disease following exposure to VOCs.

To evaluate the role of risk factors other than exposure to VOCs, other models are more appropriate. Given the absence of an association between VOC-exposure and gall bladder disease, the data were analyzed again with a conventional PROC PHGLM (no partitioning of the followup period) with TFU1 starting at age 21 (to control for age, and to restrict the analysis to adults), and SEX and DIAB (the presence of diabetes) as independent variables, ignoring the VOC variables WASH and WATER. This resulted in a larger number of cases, since no eligibility criteria were applied to TDIAG. Other variables added (one at a time) to this basic model included: cholesterol, triglycerides, kidney and liver tests, the serum levels of the pesticide DDT, PCB and PBB (polychlorinated and -brominated biphenyls), QI, alcohol use, smoking, income and education level, occupational exposure to VOCs, pesticides, or other chemicals. Statistically significant were SEX (women at higher risk), QI, DIAB, creatinine and blood urea nitrogen (both had a negative coefficient), and WORK2 (occupational exposure to solvents other than VOCs). The results of the model with the variables SEX, DIAB and WORK2, and a second model with all variables that remained significant when joined together, are shown below in Table G.6. The values of DIAB and WORK2 may reasonably be assumed to be true prior to TDIAG; this assumption is less likely valid for the other variables.

Table G.6 Results of two models, showing coefficients and p values for variables.

MODEL 1: (42 cases GALL=1&2 + 411 noncases, R=0.205, -2logL=392, model p<0.0001)

TFU = -1.563 SEX + 1.388 DIAB + 0.749 WORK2
 p = 0.0001 0.0016 0.0514

MODEL 2: (39 cases GALL=1&2 + 391 noncases, R=0.244, -2logL=351, model p<0.0001)

TFU = -1.442 SEX + 1.197 DIAB + 0.702 WORK2 + 0.055 QI - 0.1007 BUN
 p = 0.0008 0.0208 0.0726 0.0151 0.0141

G.6 Discussion

From a risk assessment viewpoint, the main problem in collecting health data is to ascertain the accuracy and the completeness of the data obtained. The data sources for determining the health status included an interview and a review of medical records. Chenodiol, used for dissolving cholesterol stones, was not among the drugs prescribed, and its use could hence not be used as a validation tool. Laboratory data cannot be used as an additional source of information for validating the diagnosis, making the validation of gall stone cases different from that for diabetes and hypertension.

The analysis was performed for 1) the case definition GALL=1, and 2) the alternative definition GALL=1&2 (which had the advantage of a larger number of eligible cases). The larger number of people in category GALL=2, however, may raise doubt as to the correctness of the interview response. Incomplete retrieval of medical records was not a problem in gathering data on diabetes, hypertension (Appendices E and F), or cancer (Chapter 8), and it is unlikely that selective deficiency in retrieving gall bladder disease would have occurred. False-positive responses may easily have originated from misinterpretation of medical information received from physicians. In any event, no significant differences were found in the results based on the two case definitions.

The negative exposure-disease associations across case definitions, analytical models, and exposure expressions, are sufficient evidence against the hypothesis of a toxic effect of VOCs. An additional support for this conclusion is the insignificance of occupational exposure to VOCs as a risk factor, as shown by the results of the conventional PROC PHGLM without VOC-related variables. The OR was increased if exposure was expressed as having lived at any time in the contaminated area (Tables G.3 and G.4). This demonstrates the danger of using an improper exposure classification. To ensure that the rejection of GALL=3 as a valid case of disease did not bias the conclusions, all analyses were repeated for a case definition which included GALL=3 as well (that is, no attempt was made to validate health data other than for TDIAG). The results did not change the conclusion of no observed adverse effect. The negative exposure-disease association is too consistent to be spurious, but no explanation was found. In any event, exposure to VOCs at the prevailing levels can be rejected as a cause of gall bladder disease.

In an attempt to identify risk factors other than VOCs, an analysis was done, using a conventional proportional hazard model, and accepting all valid cases, regardless of TDIAG and VOC exposure. As was expected from a review of the literature, SEX, QI, and DIAB were significant risk factors (3, 4, 5). No earlier reports have been found on a *negative* association of the disease with *current* levels of creatinine and blood urea nitrogen (not with uric acid). The

negative association was unexpected, since the concurrent association of gall bladder disease with diabetes would predict a positive coefficient for the kidney function tests. WORK2 as a significant risk factor was an unexpected finding. The information available was insufficient for distinguishing between a spurious and a true association between occupational exposure to solvents other than VOCs and the risk of developing gall bladder disease.

Based on the coefficients of the regression model shown in Section G.5, women are at 4.8 times higher risk of developing the disease than men, an OR which may decrease if WORK2 is replaced by other variables. The lowest OR measure was 3.2 if WORK2 was replaced by QI. Gender remained a risk factor in any of the models tested, suggesting that the difference in gender-specific IRs is not determined by either of the risk factors tested. Depending on what variables were in the model, the coefficient for DIAB ranged from 1.0 to 1.4, yielding a range of ORs from 2.7 to 4.

Diehl *et al* (6) suggested an association of diabetes and gall bladder disease, although no such association was found in the Framingham study (3). The latter suffered from a lack of controlling for confounding: the conclusions were based on simple differences in proportions. Diehl *et al* did control for obesity and age in a logistic regression model applied to a case-control study, but studied women only. Further, although obesity was included in the model, its coefficient was statistically not significant. This may have been caused by a sample not representative of the total population (40% of the population had no information on age, and was therefore not included in the analysis). It has been suggested that insulin use rather than diabetes is associated with gall bladder disease (7), which might explain the inconsistent significance of DIAB as a risk factor. The number of diabetics in the current study was too small, however, for a meaningful analysis of insulin-treated diabetes. The positive association of GALL with DIAB and QI in the current study is a complex issue. DIAB did not reach significance ($p > 0.4$) if added to a model with SEX and QI as the only other predictors. Significance occurred if QI was left out, or if another of the significant risk factors was added. QI was a significant factor in all models tested, but if this variable was replaced by OBESITY (the border between obese and nonobese is the median QI value equal to 25), both DIAB and OBESITY became insignificant in any of the combinations tested. Using a higher cutoff point for defining OBESITY (QI=30 or more), however, returned the significance of both variables. Apparently, the contradicting reports on the associations of gall bladder disease with DIAB and obesity resulted from using different definitions of obesity and, probably, diabetes.

Interestingly, despite the association of gall bladder disease with DIAB and QI, no significant association was found with cholesterol and triglycerides. The literature on the relation between gall stones and blood lipids is confusing. Virtually all studies have been based on clinical observations or autopsy findings. No study has used multivariable analysis to cope with the disturbing influences of obesity, diabetes, age, sex, fasting, and drugs affecting the cholesterol level. Stone formation (particularly of cholesterol stones), is associated with supersaturated bile. It is also dependent on nucleus formation, bile flow, the total bile acid pool, bile acid excretion, and many other known and unknown conditions controlling the growth of initiated nuclei (8). This may explain the seemingly contradictory observation that reduced cholesterolemia, resulting from a low-cholesterol diet rich in polyunsaturated fats, was positively associated with stone formation (3).

One problem in the evaluation of the association of gallstones and cholesterol is that stones are rarely pure cholesterol deposits. In the current study, no information was obtained which would allow a distinction to be made between cholesterol, pigment, and mixed type stones. Clinical testing did not include separation of the density types of lipoproteins. Cholelithiasis is usually accompanied by chronic gall bladder inflammation, at least in the stage when a patient is seeking medical help. An analysis taking account of the inflammatory component and the type of lipoprotein (while controlling for the influence of known risk factors), might bring some clarity

to the current confusing information on the role of blood lipids. In the current study, blood lipids were not a significant risk factor in any of the models tested. This may reflect a true absence of association, but may also indicate that current blood lipid levels are not a suitable indicator of past levels.

Although many studies have identified recall deficiency as a problem, no study has been found which examined the validity of the date of diagnosis, mentioned during interview, and of the diagnosis itself. Yet, in any analysis of the risk of disease following exposure, the accuracy of TDIAG is crucial for establishing the correct temporal sequence of exposure and effect. In this study, the date when exposure started was estimated through groundwater modeling (1) and from TIN (date when people moved to the study area). For estimating TDIAG, the sources of information in a retrospective study are necessarily limited to interviews and medical records. Of the 25 cases with a positive interview response and a positive medical record, one did not recall TDIAG at interview. Of the remaining 24 cases, eight had the same TDIAG, 10 people reported at interview a TDIAG earlier than that of the medical record (range 1 - 27 months, average 11 months), and six people reported a later TDIAG (range 1-37 months, average 15 months). For the analysis, the earliest of the reported or recorded dates was selected as TDIAG, with the implied risk of erroneously rejecting a correct date from a medical record if the interview TDIAG was earlier. There were no means, however, to check on the accuracy of an earlier interview TDIAG because failure to retrieve an earlier medical record could not be ruled out. All together, the data suggest that respondents were less inaccurate in recalling TDIAG for gall bladder disease than was the case for hypertension or diabetes (Appendices E and F). This may be explained by an abrupt onset of symptoms or by subsequent surgery for gall bladder disease.

In comparing interview responses with medical records, eight out of 33 people with a positive medical record responded negatively when asked about a history of gallstones. Five of the eight lived in one neighborhood and had an age range of 34 to 60 years; the other three were women of advanced age (71 - 86) spread over three other neighborhoods. For the latter three subjects, advanced age might have contributed to a recall deficiency. This does not hold for the others and (since they came from one neighborhood) their recall failure might be a common source of insufficient medical communication. The above data indicate a false-negative response rate of at least 8/33 or 24%. Too late a TDIAG was reported (at interview) at a rate of at least 6/24 or 25%. This degree of error in the health status and TDIAG is evidence of the need to extend quality control procedures to health data.

The design of the current study provided the means to directly estimate disease incidence rates. The annual IR of 1.7 to 5.7 per 1000 population of all ages (depending on the case definition and gender) has no suitable reference data. Friedman *et al* (3) reported an annual IR of 2.9 (men) and 5.9 (women) per 1000 Framingham Study population for the group age 30-62. Their case definition was comparable to GALL=1&2 and the followup duration was 10 years. The IR denominator was the number of persons at the start of the followup rather than person-time of followup. Applying this improper method for computing the IR to the Battle Creek Study, the age-adjusted IR (the Framingham age-distribution serving as the standard) for the age group 30-62 and the observation period 1976-1985 would be 6.2 for men and 7.5 for women. If the case definition is restricted to GALL=1 (positive medical record), the age-adjusted IRs are 5.4 (men) and 5.8 (women). Thus, even with a restricted case definition, the current study still has an IR for men twice that seen in the Framingham study. The lower Framingham rates might be due to the fact that no records were obtained from physician's offices outside hospitals.

The National Health Interview Surveys (NHIS) provide national estimates of the prevalence of many diseases on a continuous basis. In theory, the results of these surveys should be preferred over that of the Framingham study, as all age groups are included. The average NHIS gall bladder disease prevalence rates reported for the period 1979-1981 are 0.3% for males and 1.0% for females (9). These rates are one order of magnitude lower than those observed in this

study or in the Framingham study (the latter for the age group 30-62). Although NHIS study populations are restricted to samples of the non-institutionalized population, the most likely explanation for the very low rates is that NHIS studies entirely rely upon interview responses, since review of medical records is not a part of the study design. At any rate, the evidence is that the efficacy of case-finding in the Battle Creek Health Study compares very favorably with the results of other longitudinal or cross-sectional studies, regardless of how a case is defined. It may therefore be inferred that failure to identify cases has not played a role in the outcomes of the statistical analysis of the association between gall bladder disease and exposure to VOCs.

G.7 Conclusions and Summary

Following the detection of groundwater contaminated with chlorinated volatile organic compounds (VOCs), a retrospective cohort study was conducted involving 749 people with an average followup period of approximately 13 years, and a 1:2 ratio of exposed to unexposed subjects. The study followed the principles of risk assessment for toxic chemicals. It was not designed to specifically study gall bladder disease, but the presence of 55 potential cases of gall bladder disease and a possible excess of cases in the contaminated area justified a detailed analysis. With a case defined as any person with either a positive medical record or a positive interview response indicating hospitalization (albeit no medical record could be retrieved), the validation procedure yielded 46 valid cases of gall stones and/or chronic inflammation of the gall bladder. This number results in a prevalence rate as of July 1985 of 6.3% (males 3.2% and females 8.9%), and an annual incidence rate for the period 1970-1985 of 3.89 per 1000 population of all ages (males 1.87 and females 5.67). The reference data for the general population, compiled by the National Center for Health Statistics, and an early study on a selected population, indicate that these current rates are a multiple of what was observed elsewhere. This may reflect the incomplete data base of the earlier studies, but some true excess of cases of gall bladder disease in the current study population cannot be ruled out. Using univariable and multivariable analysis, the following conclusions about the risk of acquiring the disease can be drawn from the study:

- 1) Exposure to VOCs in drinking water is *not* related to the risk of developing gall bladder disease.
- 2) The major risk factors for developing the disease are sex (women are at 3 to 5 times higher risk than men), diabetes (odds ratio 2.7 to 4), and obesity. The association with diabetes and obesity is a complex issue, probably due to mixing several types of diabetes and to the definition of obesity. The study provided information on the hitherto unknown inverse relation of gall bladder disease and blood urea and creatinine, and on the incidence rate of gall bladder disease in the total population. Incidence and prevalence were much higher than observed in earlier studies of other populations.
- 3) DDT, PCB, and PBB (at the prevailing serum levels), education and income level, occupational exposure to a variety of chemicals, smoking, alcohol, serum lipids, liver enzymes, and uric acid did not show an association with the occurrence of gall bladder disease.

This study appears to be the first epidemiologic study of gall bladder disease, involving a sample of the general population *and* using multivariable analysis. Other important findings from the study are: a) any historical survey for medical data which is limited to an interview is bound to seriously underestimate the rates of disease occurrence; b) people are remarkably inaccurate in recalling the occurrence of gall bladder disease and the date of diagnosis.

G.8 References

1. Freni S. C. and Phillips, D. L. Estimation of the time component in the movement of chemicals in contaminated groundwater. *Environ. Health Perspect.* 74: 211-221 (1987).
2. Harrell, F. E., Jr. The PHGLM procedure. In: *SUGI Supplemental Users Guide*, SAS Institute Inc., Cary NC, 5th edition 1986: 438-466.
3. Friedman, G. D., Kannel, W. B., and Dawber, T. R. The epidemiology of gall bladder disease: observations in the Framingham study. *J. Chron. Dis.* 19: 272-292 (1966).
4. Bennion, L. J. and Grundy, S. M. Risk factors for the development of cholelithiasis in man (second of two parts). *N. Engl. J. Med.* 299: 1221-1227 (1978).
5. Kern, Jr., F. Epidemiology and natural history of gallstones. *Seminars Liver Dis.* 3: 87-96 (1983).
6. Diehl, A. K., Gaskill, S. P., Hazuda, H. P., et al. Prevalence of clinical gall bladder disease in Mexican-American, Anglo and Black women. *South Med. J.* 73: 438-443 (1980).
7. Capron, J. P., Dupas, J. L., and Capron-Chivrac, D. Les facteurs associes a la lithiase biliaire cholesterolique. *Gastroenterol. Clin. Biol.* 4: 63-77 (1980).
8. Bennion, L. J. and Grundy, S. M. Risk factors for the development of cholelithiasis in man (first of two parts). *N. Engl. J. Med.* 299: 1161-1167 (1978).
9. National Center for Health Statistics. *Prevalence of Selected Chronic Conditions - United States, 1979-81. Data from the National Health Survey Series 10, No. 155.* U.S. Department of Health and Human Services, DHHS Publication no. (PHS) 86-1583, Hyattsville, MD, July 1986.

Appendix H

Reproductive Effects of Exposure to Volatile Organic Chemicals

**Stan C. Freni, MD, PhD, DrPH
Riduan Joesoef, MD, PhD
Debbi L. Kotlovker**

**Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

H.1 Introduction

In 1981, the aquifer at the Verona Well Field, Battle Creek, Calhoun County, Michigan, was found to be contaminated with chlorinated, short chain, aliphatic volatile organic chemicals (VOCs). Most of the city wells, and a large number of residential wells were affected. Subsequently, VOCs were also detected in the water of a number of private wells in Springfield, which is adjacent to Battle Creek, and in Dowagiac, Cass County, Michigan. In 1984, Phase I of a two-phase epidemiologic study was initiated to collect demographic data to form exposed and reference cohort populations for Phase II, a retrospective followup study, which started in July 1985. In Phase II, information was collected on morbidity and exposure for the above study cohorts. This study involved 251 current and past residents from exposed and 498 from comparison neighborhoods. The groups were matched for sex, age, family size, and neighborhood. Since 94% of the participants were white, race was ignored in the analysis.

Since the study focused on the potential effects of a predefined exposure, the principles of risk assessment for chemicals were applied. These included:

- 1) Exposure must be quantified as a dose and in time.
- 2) A dose-response effect must be determined.
- 3) A disease is attributable to exposure only if exposure has preceded the disease.
- 4) Quality control procedures must be extended to the quality of health data.
- 5) Uncertainties in the data and methods, and their effect on the study results, must be properly addressed.

Toxicologic information indicates that VOCs may have toxic effects on fetuses of pregnant laboratory animals exposed to high doses. These effects include fetal absorption, fetal death, smaller size, and delayed fetal development, which may be translated to humans in terms of miscarriage, fetal death from causes other than labor, prematurity, or low birth weight. This Appendix reports on the analysis of these pregnancy outcomes in association with exposure to VOCs in drinking water. Birth defects may be another expression of fetal toxicity. The results of the analysis of birth defects are given in Chapter 8.

H.2 Health Assessment and Case Definition

All hospitals serving Calhoun and Cass counties, and other regional hospitals in and around these counties (see Appendix B) were canvassed in a survey for medical records. In addition, the offices of physicians and hospitals mentioned by the study participants were contacted to retrieve records. Copies of all death records were obtained. The records and certificates were abstracted for diseases occurring as of 1970.

An in-person interview and a limited clinical examination were conducted. The questionnaire asked women to list all pregnancies, to indicate the date and nature of each pregnancy outcome (full term alive, premature alive, stillborn, miscarriage, birth defect, birth weight), and to provide the names and addresses of the hospitals involved. In validating the interview responses, the medical records, if available, were considered the correct reference. Since the abstracting of medical records was not oriented specifically to pregnancy outcomes, a number of records on these events could have remained unnoticed by the abstractors. Often, the event was found noted in the file of the child rather than the mother. Further, not all records were retrievable

or complete, particularly those from private physician's offices. Categorization of pregnancy outcomes was based on the following definitions:

MISCARRIAGE:

Any event reported as such by the respondent or in the medical record. In case of discrepancy, the latter prevailed. Recognizing the known unreliability of responses regarding induced abortions, the questionnaire purposely addressed "miscarriage or spontaneous abortion". Most miscarriages were not covered by a medical record. In only one case of miscarriage was the duration of the pregnancy mentioned in a record: 18 weeks.

STILLBIRTH:

Some respondents considered a stillborn fetus a miscarriage, even if the birth weight (BW) was more than 453 g and the medical record mentioned premature stillbirth. Except for these cases, there were no discrepancies with regard to STILLBIRTH. Cases of stillbirth are usually also premature cases on the basis of either the medical record or the weight limits described above.

PREMATURITY:

This can only be diagnosed with certainty if the menstrual history and certain physiological and physical features of the newborn are known. In the absence of these data or of a medical record, validation of interview responses was necessarily based on rather arbitrary criteria. Since there were no medical record cases of prematurity with BW more than 2240 g, newborns reported by the interviewee as premature were considered mature if the BW exceeded 2500 g. In 12 such cases, the BW ranged from 2552 - 3459 g. The lowest BW of a full term newborn found in a medical record was 2466 g. Therefore, newborns under 2000 g reported as mature by the interviewee (two cases with a BW of 1928 and 1985 g) were classified as premature. The limit of 2000 g (in effect if no medical record of the pregnancy outcome was available) takes into account the fact that the lower BW limit of medical record supported cases of mature newborns may be too high because of small numbers. It was also assumed that the BW reported by the mothers was too low if the number of ounces was not recalled.

LOW BW and VERY LOW BW:

Reproductive outcomes were also categorized in terms of LOW BW (lower than 2500 g) and VERY LOW BW (lower than 1500 g). These limits correspond with the common clinical use of the terms, and are probably of limited value in view of the clear inclination of respondents to express weights in pounds only, ignoring ounces, while often no birth weight could be recalled. Unfortunately, BW was *not* among the data abstracted from medical records.

TDIAG:

The dates of the pregnancy outcomes reported by interviewees were compared with independently collected information on birth dates of members of the household (if the children participated in the study), and with dates found in medical records. As the latter usually referred to discharge dates, differences of 1 month were ignored. In case of discrepancy, birth dates were given priority over dates from interview responses or medical records. For the exposure-effect analysis, all pregnancy outcomes were eligible if TDIAG was later than 1-1-1970 and later than the date of moving to the study area (TIN).

H.3 Exposure Assessment

The individual "total accumulated exposure" (TAE) level of participants in the Verona area, accumulated during the period of residency, was estimated as the area under the $\log C = T$ curve described in Appendix C and elsewhere (7). In this notation, $\log C$ is the natural log of the concentration C of a given chemical in the well water sample (in parts per billion or ppb), and T is the date of sampling expressed in months since 1-1-1970. TAEVOC is the sum of chemical specific TAEs in units of ppb-months, corrected for TIN and the date a person stopped drinking the contaminated water (TSTOP). For cases, TAEVOC was computed until TDIAG if this was earlier than TSTOP. An alternative exposure expression is TAEVCL, equal to TAEVOC + the TAE for chloroform in city water. Other exposure expressions were 1) DOSEVOC and DOSEVCL, equal to the product of WATER (amount of unheated tap water consumed per day at home) with TAEVOC and TAEVCL, and 2) having lived in the contaminated area. These multiple exposure expressions offer an opportunity for performing exposure-effect analyses from different viewpoints. However, DOSEVCL is the option most consistent with risk assessment principles and toxicology. In the analysis, exposure variables were used as dichotomous variables, variables with the values zero-low-high (the boundary between low and high is determined by the rounded median of all positive values), or as continuous variables.

As explained in Chapter 4, secondary routes of exposure to VOCs are addressed by inserting WASH (frequency and duration of showering and bathing, Equation 7 in Chapter 4) as a covariate in the logistic regression model. In models containing TAEVOC or TAEVCL, WATER and WASH were added as a covariate. In models with DOSEVOC or DOSEVCL, WATER was already included in the exposure variable, and WASH was replaced by WASHVOC or WASHVCL. WASHVOC is WASH times the sum of VOC concentrations (CS) in the current water sample, and WASHVCL is WASH times the sum of CS + 6 (chloroform concentration for residents on city tap water).

As described in Chapter 6, the exposure status was determined for each pregnancy separately. That is, the exposure was calculated as TAE or DOSE incurred in the period from pregnancy outcome N to pregnancy outcome N-1. The toxic effect of a chemical on the fetus has the characteristics of an acute or subacute effect of exposure rather than a chronic effect. Thus, if exposure stopped before a pregnancy was initiated, the exposure status would equal zero, unless the chemical is persistent in the human body. VOCs are rapidly removed from the body, however, within days to a few weeks depending on the exposure levels. In agreement with this line of reasoning, it was decided that a woman would be considered unexposed for the index pregnancy, if exposure stopped more than 2 months prior to conception. (Dates of conception were approximated as 9 months before a term birth, 4 months before a miscarriage, and 7 months before a premature birth). These 2 months should not be viewed as an accurate estimate of the period required for a toxic effect to fade out after cessation of exposure. It is rather a recognition that, with respect to reproductive events, a woman does not remain exposed indefinitely because of a past exposure.

H.4 Statistical Analysis

The analysis included 1) a univariable analysis, involving fourfold tables with the odds ratio (OR) as the measure of association (computed as described in Chapter 6), and 2) a multivariable logistic regression model (8). As most women had two or more pregnancies, the unit of observation was the pregnancy outcome, and not the study participant. In the univariable analysis, testing for statistical significance was done with a one-sided Fisher's exact test. A conventional logistic regression analysis of a data set of pregnancy outcomes may yield biased results, as the outcome of a pregnancy is influenced by the outcome of previous pregnancies. Therefore, an alternative analysis was performed using a modification proposed by Liang and

Zegers (9), which corrects the coefficients of the conventional model for possible interdependency among the observations. The basic multivariable model contained the independent variables TAE or DOSE, WATER, WASH, and AGE at the outcome of the pregnancy. The number of eligible cases was too small to include more potential confounding factors. Prematurity was not included in this analysis because of the very small number of eligible cases (N = 7).

H.5 Results

Of the 400 women participating in Phase II of the Battle Creek Health Study, 316 were age 15 and over, and 249 of these had had at least one pregnancy. There were no pregnancies reported by younger participants. As the abstractors of the medical records did not focus on reproductive outcomes, the lack of a record does not necessarily imply a failure to retrieve a record, or that the interview response was incorrect. There were no medical records of "death due to complications of labor", but it is possible that some of the stillbirths represent such cases. Evaluation of the 207 pregnancy events covered by a medical record provided the following summary of discrepancies between the sources with regard to abnormal pregnancy outcomes:

Table H.1 Discrepancies between data sources with respect to abnormal pregnancy outcomes.

Interview	Medical record
1 case of unknown outcome	1 normal live birth
17 cases of prematurity	14 premature, 2 mature newborn
	1 no information on maturity
9 cases of miscarriage	7 miscarriages, 1 stillbirth,
	1 breakthrough bleeding postpartum
8 cases of still birth	6 stillbirths, 2 miscarriages
7 pregnancies <i>not</i> reported	2 normal live births, 5 miscarriages

As stated earlier, the medical records were accepted as correct in all cases of discrepancy, except for two in which the medical record was clearly wrong. The first case was a report of threatened abortion 2 months before a mature child was born. The second case was a report of a premature birth of 35 weeks, although the child's birth weight was 4281 g and the mother had reported a full-term pregnancy outcome.

The 249 women who were ever pregnant reported a total of 802 pregnancies, and seven additional pregnancies were found in the medical records. In one case, the interviewee reported a miscarriage, but the medical record showed that this was a case of breakthrough bleeding 2 months post partum, and not an aborted new pregnancy. The remaining 808 events included: 654 mature live births (81%), 88 miscarriages (11%), 13 stillbirths (2%), 101 cases of "pregnancy loss" (stillbirth or miscarriage, 13%), 29 premature births (4%), 59 cases of low BW (7%), and 11 cases of very low BW (1%). As a given event may fit in *more than one* category (e.g., stillbirth at week 28 with a very low BW), numbers and percentages *add up to more than 808 or 100%*. Table H.2 depicts the age-specific prevalence for the various types of abnormal events observed among the 312 women alive as of July 1985. Of the 249 women ever pregnant, 100 had at least one abnormal outcome, equal to a crude prevalence for any abnormal outcome of 40.2%. The National Center for Health Statistics (NCHS) estimated that the prevalence of pregnancy loss (miscarriage plus stillbirth) in the U.S. population was 26.0 per 100 women age 15-44 ever pregnant (10). Using the NCHS age-distribution as the standard, the age-adjusted rate of pregnancy loss in the current study (for the population age 15-44) was 28.5/100 women ever pregnant and alive as of

July 1985. The 95% confidence limits (δ) are 21.6-35.9%. Thus, the prevalence of women with at least one miscarriage or stillbirth in this study is slightly higher than the national average, but the difference is not statistically significant. The age-adjusted prevalence rate for the other pregnancy outcomes was: miscarriage 25.4%, stillbirth 4.2%, prematurity 10.9%, low birth weight 16.2%, very low birth weight 4.5%, any abnormal outcome 33.4%.

Table H.2 Prevalence of reproductive events among women age 15+, alive as of July 1985.
BW = birth weight.

Age	Number of women	Number ever pregnant	Prevalence/100 women ever pregnant, by age-category					
			Miscarriage	Still-birth	Pregnancy loss	Prematurity	Low BW	Very low BW
15-19	39	5	40.0	0	40.0	0	0	0
20-24	38	21	14.3	0	14.3	4.8	9.5	4.8
25-29	41	39	38.5	5.1	38.5	12.8	17.9	0
30-34	29	27	22.2	7.4	29.6	11.1	14.8	7.4
35-39	36	36	25.0	2.8	27.8	8.3	13.9	2.8
40-44	26	21	19.0	4.8	23.8	19.0	28.6	9.5
45-49	18	17	17.6	5.9	17.6	11.8	17.6	5.9
50-59	33	32	35.0	3.1	28.1	9.4	31.3	6.3
60-69	35	35	20.0	5.7	25.7	14.3	28.6	5.7
70 +	17	16	18.8	6.3	18.8	0	6.3	0
15-44		149	26.2	4.0	28.9	10.7	16.1	4.0
Total		249	24.1	4.4	26.9	10.4	19.3	4.4
			Prevalence/100 women total population (ever or never pregnant)					
15-44	209		18.7	2.9	20.6	7.7	11.4	2.9
Total	312		19.2	3.5	21.5	8.3	15.4	3.5

Only 199 pregnancy outcomes were eventually found eligible for analysis; that is, they occurred later than 1-1-1970 and later than TIN. For exposure defined as having lived in the contaminated area, the only eligibility requirement is that TDIAG > TIN, since the criterion of TDIAG later than 1-1-1970 is irrelevant because this exposure definition is unrelated to actual exposure. When this exposure definition is used, the number of events is thus larger than for other exposure expressions. Table H.3 depicts ORs for the different pregnancy outcomes and the various expressions of exposure. The ORs are above unity only if exposure is expressed as having lived in the contaminated area. The more meaningful exposure expressions TAE and DOSE are associated with ORs below unity, equivalent to an excess of abnormal events in the reference population. None of the analyses, however, yielded an OR with significant p-values. The dose-response relationships were inconsistent: they were negative for TAEVOC and TAEVCL, and mixed for DOSEVOC and DOSEVCL. This likely reflects a chance effect, as the ORs consistently lower than unity indicate absence of an effect of exposure. Further, in none of the dose-response tables (in which the exposure was stratified) was the p-value lower than 0.1.

Table H.3 Summary of the results from fourfold tables of 199 pregnancy outcomes eligible for analysis. The results are shown as the odds ratio (OR) and the p-value from a one-sided Fisher's exact test. Because for "Exposure = Contaminated AREA" the only eligibility criterion applied was that TDIAG must be later than TIN, the number of abnormal events is larger than for the other exposure expressions.

Pregnancy outcome	Contaminated AREA		Exposure expression							
	OR	p	TAEVOC		TAEVCL		DOSEVOC		DOSEVCL	
			OR	p	OR	p	OR	p	OR	p
Miscarriage	1.45	0.24	0.98	0.61	0.82	0.41	0.87	0.56	0.77	0.36
	N = 27		N = 23		N = 23		N = 23		N = 23	
Pregnancy loss	1.20	0.56	0.88	0.54	0.68	0.25	0.78	0.49	0.65	0.22
	N = 30		N = 25		N = 25		N = 25		N = 25	
Prematurity	0.50	0.30	0	0.25	1.23	0.55	0	0.33	1.39	0.48
	N = 10		N = 7		N = 7		N = 7		N = 7	
Low birth weight	1.38	0.37	0.45	0.39	0.72	0.41	0.55	0.49	0.81	0.49
	N = 15		N = 11		N = 11		N = 11		N = 11	
Any abnormal event	1.29	0.28	0.63	0.29	0.77	0.32	0.56	0.27	0.77	0.31
	N = 44		N = 32		N = 32		N = 32		N = 32	

All coefficients of the exposure variables in the conventional logistic model were negative, that is, ORs computed from these coefficients would be below unity, and the dose-response would be negative. The coefficients had p-values larger than 0.5, the model p-value was usually larger than 0.3, and the model R-statistic of the models did not exceed 0.06. This is another indication that exposure to VOCs is not the cause of the difference in the rate of occurrence of abnormal pregnancy outcomes between the exposed and unexposed cohorts. The results of the modified logistic regression were little different with regard to the direction of the coefficients and the lack of statistical significance.

H.6 Discussion

The analysis presented does not assume complete recall on the part of the interviewees or total accuracy of the data. It does assume that misclassification due to erroneous interview data is nondifferential as to the exposure status. The comparison of interview data with that from medical records showed a much closer agreement between sources than observed for health events other than pregnancy outcomes. There was no evidence that inaccuracies in the data were related to the exposure status. It is obvious that no estimate can be made of failure to recall miscarriages not covered by a medical record. Such recall deficiency might be more prevalent among older women. However, as the age-distributions of the cohorts are quite comparable (Chapter 2), this potential inaccuracy in the data is unlikely to have affected the outcome of the analysis.

The age-adjusted prevalence for any "abnormal outcome" is 33.4%, in other words, one third of women (participating in the study and alive as of July 1985) experienced an abnormal outcome

at least once. No reference rates for any abnormal outcome were found. Compared to the NCHS data for the U.S. population (10), the prevalence of miscarriage or stillbirth in the study population is slightly higher, which suggests that the case-finding survey has not suffered from incompleteness. The magnitude of the true rate of early spontaneous fetal loss is unknown. Erhardt estimated that 4 to 8% of the pregnancies in New York City might end before week 11 of gestation (4), a percentage far below the 28% (age-adjusted) fetal loss in the current study, which was not designed to identify very early losses unnoticed by the women or their physician. Erhardt's low estimate is almost certainly the result of incomplete case retrieval. The extent of very early fetal loss can be adequately estimated only in a prospective study with a full medical evaluation (including hormone assays) of all delayed "menstruations". Whether these very early losses are an issue in the toxicology of chemicals is not certain, however, since hormonal disturbances or genetic defects are probably the major causes of this type of early fetal loss.

VOCs are known to be fetotoxic in animal studies at very high doses. At lower doses, Schwetz and coworkers (12) failed to detect adverse reproductive effects in a study on rats and mice exposed during days 6-15 of gestation to 305 (TCE, PCE) and 872 (TCA) ppm in air for 7 hours a day. The only possible abnormal finding was a slight increase in the incidence of fetal absorptions in rats treated with PCE. This suggests that there might be a non-zero threshold of exposure, below which there is no risk of adverse effects. If the exposure levels from the study by Schwetz constitute the margin between NOEL and LOEL (No and Lowest Observed Effect Level), the actual NOEL may be somewhat lower, say, 250 ppm in air. This equals approximately 5.1 g/day (average conversion factor of 6.0 for VOC concentration to weight, average respiratory uptake of 60%, and conversion from 8 hours/workday to continuous exposure), or about 2500 ppm in water. This is 5 to 6 orders of magnitude higher than the prevailing exposure levels in the Battle Creek Health Study.

The results from univariable and multivariable analyses unequivocally indicate that, at the prevailing levels of water contamination, VOCs had no adverse effect on pregnancy outcomes. The only OR (slightly) above unity was for TAEVCL and DOSEVCL in relation to prematurity. This is of no importance (as is reflected in the very high p-value) because of the very small number of cases. The shift of just one case from exposed to unexposed would have reversed the OR to well below unity. Also, the OR for TAEVOC and DOSEVOC was zero (no cases in the exposed cohort, if exposure to chloroform is ignored). The absence of a positive exposure-effect association remained if the analysis was restricted to the Verona exposed and the Calhoun county reference cohort (for reason of better exposure data) or if birth defects were the pregnancy outcome of interest (Chapter 8). The inappropriateness of an analysis that does not adhere to basic risk assessment principles is illustrated by the fact that the ORs were above unity (statistically insignificant in spite of a large number of cases) if exposure was expressed as having ever lived in the contaminated area (disregarding the temporal sequence of exposure-effect and the actual exposure status).

The maximum TAE level preceding any given pregnancy was 1,523,594 ppb-months TAEVOC, the second highest level was 409,398 ppb-months, and only 11 women had more than 10,000 ppb-months at the time a pregnancy terminated. The maximum level of 1.5 million ppb-months, if concentrated within a period of 1 year, is equivalent to an ingested dose of 31 mg/day for 1 year. (This case had a reported daily consumption of one 8 oz. glass of unheated tap water.) If the reported amount of water consumed was erroneous, and assuming a true consumption of 1 liter/day, the ingested dose would have been 127 mg/day. The woman with this dose had a pregnancy with a normal outcome and no previous pregnancies. She had a second pregnancy 1 year after exposure stopped, and the outcome was also normal. This observed human NOEL of 0.127 g/day is many orders of magnitude lower than the NOEL of 5.1 g/d observed in the animal reproductive study discussed above.

If exposure to VOCs is ignored because of the absence of evidence of a toxic effect, the entire data set of 808 pregnancies may be analyzed for other risk factors. Typical risk factors for abnormal pregnancy outcomes are a positive history of previous abnormalities and smoking. Strobino and coworkers (13) addressed the issue of repeated miscarriage by comparing "repeaters" (more than three miscarriages) with "sporadics" (one miscarriage out of at least three pregnancies) and "normals" (no miscarriage out of at least three pregnancies). Prominent risk factors for recurrent miscarriage were: a delay in the conception period of the index pregnancy, cervical incompetence, and a history of low birth weight or premature deliveries. In the current study, an exploratory analysis, using a logistic regression model, showed that the only significant risk factor for miscarriage was a previous history of miscarriage. Prematurity was associated with smoking and occupational exposure to solvents other than VOCs. No risk factors were found for low birth weight. An in-depth analysis may substantially change these findings, however.

As shown by the ORs for exposure defined as having lived in the contaminated AREA, there is a surplus of abnormal pregnancy outcomes in the contaminated areas unrelated to VOCs. Whether or not this can be explained by differences in the risk factors found in the exploratory analysis was not investigated. Adverse reproductive effects from environmental exposure have been reported for lead and arsenic (11), and polychlorinated biphenyls or PCBs (5). However, these studies did not show a clear dose-responsiveness; the analyses were either an analysis of variance or a testing for differences in means, both of which are inadequate tools for analyzing exposure-effect associations. In the current study, no associations were observed with PCB, PBB (polybrominated biphenyls) or the pesticide DDT (no information was available on arsenic levels in the body). Much more time would be needed for a detailed analysis of risk factors other than VOCs. As indicated repeatedly in the report, such analysis was not the primary goal of the study.

H.7 References

1. Bierman, J. M., Siegel, E., French, F. E., and Simonian, K. Analysis of the outcome of all pregnancies in a community, the Kauai pregnancy study. *Am. J. Obst. Gynec.* 91: 37-45 (1965).
2. California Department of Health Services (CDHS). *Pregnancy Outcomes in Santa Clara County 1980-1985*. CDHS, Berkeley, CA, May 1988.
3. California Department of Health Services (CDHS). *Water Exposure and Pregnancy Outcomes*. CDHS, Berkeley, CA, May 1988.
4. Erhardt, C. E. Pregnancy losses in New York City, 1960. *Am. J. Publ. Health* 53: 1337-1352 (1963).
5. Fein, G. G., Jacobson, J. L., Jacobson, S. W., et al. Prenatal exposure to polychlorinated biphenyls: effects on birth size and gestational age. *J. Pediat.* 105: 315-320 (1984).
6. Fleisch, J. L. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, 2nd edition, New York, 1981: p. 4.
7. Freni, S. C. and Phillips, D. L. Estimation of the time component in the movement of chemicals in contaminated groundwater. *Environ. Health. Perspect.* 74: 211-221 (1987).
8. Harrell, F. E., Jr. The LOGIST procedure. In: *SUGI Supplemental Users Guide*, SAS Institute Inc., Cary NC, 5th edition 1986: 269-293.
9. Liang, K. and Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika* 73: 13-22 (1986).
10. National Center for Health Statistics. *Fecundity, Infertility, and Reproductive Health in the United States 1982*. Data from the National Survey of Family Growth, series 23, No. 14. Hyattsville, MD, 1987.
11. Nordstrom, S., Beckman, L., and Nordenson, I. Occupational and environmental risks in and around a smelter in northern Sweden - III. Frequencies of spontaneous abortions. *Hereditas* 88: 51-54 (1978).
12. Schwetz, B. A., Leong, B. K. J., and Gehring, P. J. The effect of maternally inhaled perchloroethylene, methylchloroform, and methylene chloride on embryonal and fetal development in mice and rats. *Toxicol. Appl. Pharmacol.* 32: 84-96 (1975).
13. Strobino, B., Fox, H. E., Kline, J., et al. Characteristics of women with recurrent spontaneous abortions and women with favorable reproductive histories. *Am. J. Publ. Health* 76: 986-991 (1986).

Appendix I

**Algorithm and Computer Programs for
Case Eligibility and Estimation of TAEVOC and TAEVCL**

**Stan C. Freni, MD, PhD, DrPH
Center for Environmental Health and Injury Control
Centers for Disease Control
Atlanta, Georgia**

JOB CONTROL LANGUAGE

```
//SCF1 JOB NC00,FRENI,CLASS=F,TIME=1,MSGCLASS=2,MSGLEVEL=(2,0)
// EXEC SAS,OPTIONS=MACRO,WTR=2
//IN1 DD DSN=SCF1.VERONA3,DISP=SHR
//IN2 DD DSN=SCF1.EPIFILE2,DISP=SHR
//OUT DD DSN=SCF1.DIABFILE,DISP=(NEW,CATLG),UNIT=FILE,SPACE=(TRK,(15,20))
```

*** VERONA3 is a disk data set with the exposure part of the questionnaire, and EPIFILE2 is a disk data set containing the most important data on exposure, laboratory results, water contamination levels, and edited data on occupational exposure, water consumption (WATER), and water use for bathing and showering (WASH). DIABFILE becomes the disk data set for the analysis of diabetes with exposure.

PROGRAM FOR APPLYING MACRO FOR TAE; AND ALGORITHM FOR CASE DEFINITION

```
*** This version is for EXPOSURE TO END OF INDEX-PERIOD. For the version of
EXPOSURE TO END OF PRECEDING PERIOD, change macro statements in lines 21 and
33 for T from T=12*N into T=12*(N-1), and do as described in lines 59a and 60,
77a, and 99a 99b.

001 %MACRO VOCS;
002 %MACRO TAEMAC(V);
003 IF GROUP=1 & &V=1 THEN T&V=TS; *** GROUP=1 is Verona exposed ***;
004 IF &V > 1 THEN DO; L&V=LOG(&V); T&V=ROUND(TS-(L&V/B&V)); END;
005 IF (GROUP=5 OR GROUP=7) & &V > 1 THEN T&V=0;
*** T&V represents TTCA, TPCE, etc., which is T1 for TCA, PCE, etc.
*** GROUP=5 is Dowagiac, GROUP=7 is Springfield cohort.
*** TTCA TPCE etc. is set at zero for Dowagiac and Springfield ***;
006 IF GROUP=1 & &V>0 & TIN < T&V & TSTOP GE T&V
007 THEN TAE&V = ROUND((EXP(B&V*(TSTOP-T&V)) - 1)/B&V);
*** B&V represents b1 (slope) for TCA PCE etc. ***;
*** TAE&V represents TAE (area under CT-curve) for TCA, PCE, etc. ***;
008 IF GROUP=1 & &V>0 & TIN GE T&V & (TSTOP-TIN) GE 0
009 THEN TAE&V=ROUND((EXP(B&V*(TSTOP-T&V)) - EXP(B&V*(TIN-T&V)))/B&V);
010 %MEND TAEMAC;
011 %TAEMAC(TCA) %TAEMAC(PCE) %TAEMAC(TCE) %TAEMAC(DCE)
012 %TAEMAC(DCA) %TAEMAC(CIS) %TAEMAC(EDC);
013 TAEVOC=SUM(TAETCA,TAEPCE,TAETCE,TAEDCE,TAEDCA,TAECIS,TAEEDC);
014 IF GROUP=5 OR GROUP=7 THEN TAEVOC1 = CS * (TSTOP - TIN);
*** CS = sum of VOC-specific concentrations in current water sample ***;
015 IF EXPOSE=1 THEN TAECHL=CHL*(TOUT-TSTOP); *** CHL = chloroform ***;
016 IF GROUP=2 THEN TAECHL=CHL*(TOUT-TIN);
*** GROUP=2 is Battle Creek city cohort ***;
017 TAEVCL=SUM(TAEVOC,TAECHL);
018 IF TAEVOC=. THEN TAEVOC=0; IF TAEVCL=. THEN TAEVCL=0;
019 %MEND VOCS;
020 %MACRO PERIOD1(N); *** MACRO for 1970-1976, excludes Verona exposed ***;
021 T = 12*N; IF (GROUP=2 OR GROUP=5 OR GROUP=7) & TIN LE T;
022 IF TSTOP>T THEN TSTOP=T; IF TOUT>T THEN TOUT=T;
*** TIN = date a person moved into study area ***;
*** TOUT = date a person left study area
*** TSTOP = date a person stopped drinking contaminated water,
023 %MACRO ONE(V);
```

```

024 IF (GROUP=5 OR GROUP=7) THEN TAE&V = &V * (TSTOP - TIN);
025 %MEND ONE;
026 %ONE(TCA) %ONE(PCE) %ONE(TCE) %ONE(DCE) %ONE(DCA) %ONE(CIS) %ONE(EDC);
027 IF (GROUP=5 OR GROUP=7) THEN TAECHL = CHL * (TOUT - TSTOP);
028 IF GROUP=2 THEN TAECHL = CHL * (TOUT - TIN);
029 VOC&N = SUM(TAETCA,TAEPCE,TAEDCE,TAETCE,TAEDCA,TAECIS,TAEDDC);
030 VCL&N = SUM(VOC&N,TAECHL);      KEEP ID VOC&N VCL&N;
*** VOC&N is VOC1, VOC2, .... VOC16, which are the period-specific TAEVOC
*** computed from T1 (or TIN, whichever is later) to end of index period;
*** VCL&N represents the same for period-specific TAEVCL ***;
031 %MEND PERIOD1;
032 %MACRO PERIOD2(N);
*** this MACRO is for the period 1977-1985 for all exposed cohorts
(EXPOSE=1) and for Battle Creek city cohort (GROUP=2) ***;
033 T = 12*&N; IF (GROUP=2 OR EXPOSE=1) & TIN LE T;
034 IF TOUT>T THEN TOUT=T; IF TSTOP>T THEN TSTOP=T;
035 %MACRO TWO(V);
036 IF GROUP=1 & &V>0 & TIN<T&V & TSTOP GE T&V
037 THEN TAE&V = ROUND((EXP(B&V*(TSTOP-T&V)) -1)/B&V);
038 IF GROUP=1 & &V>0 & TIN GE T&V & (TSTOP-TIN) GE 0 THEN
039 TAE&V=ROUND((EXP(B&V*(TSTOP-T&V)) - EXP(B&V*(TIN-T&V)))/B&V);
040 IF GROUP=5 OR GROUP=7 THEN TAE&V = &V * (TSTOP-TIN);
041 %MEND TWO;
042 %TWO(TCA) %TWO(PCE) %TWO(TCE) %TWO(DCE) %TWO(DCA) %TWO(CIS) %TWO(EDC);
043 IF GROUP NE 2 THEN TAECHL = (TOUT-TSTOP) * CHL;
044 IF GROUP=2 THEN TAECHL = CHL * (TOUT - TIN);
045 VOC&N = SUM(TAETCA,TAEPCE,TAEDCE,TAETCE,TAEDCA,TAECIS,TAEDDC);
046 VCL&N = SUM(VOC&N,TAECHL);      KEEP ID VOC&N VCL&N;
047 %MEND PERIOD2;

```

APPLICATION OF MACRO FOR TAE; ALGORITHM FOR CASE DEFINITION

```

048 DATA RAW1; SET IN1.VERONA3 (KEEP=ID YROUT YRSTOP);
049 DATA RAW2; SET IN2.EPIFILE2;
050 CS = SUM(TCA,PCE,TCE,DCE,DCA,CIS,EDC);
051 KEEP ID SEX AGE TCA PCE TCE DCE DCA CIS EDC CS TS GLUC HBA1C GROUP
052 YRIN TIN TSTOP TOUT TQUEST MOBIRTH YRBIRTH TDEATH EXPOSE;
053 DATA RAW; MERGE RAW1 RAW2; BY ID;
054 DATA DISEASE; INFILE LISTING;
* This example is for diabetes, but the program is applicable in general.
* This data set contains all information on DIABETES, derived from the
* questionnaire responses (META1), the medical record abstracts (P250),
* list of medications, laboratory file, and questionnaire responses on
* insulin allergy. The following are the diagnostic codes:
* DIAB=1 IF (HOSP=1 or QUEST=1) & (MED=1 OR GLUC/HBA1C high
* DIAB=2 IF BOTH HOSP=1 & QUEST=1, WHILE MED=. AND GLUC/HBA1C=normal
* DIAB=3: NO RECORD, NO RESPONSE, GLUC>121 & HBA1C>9.2 (>110% of NORM
* DIAB=4: NO RECORD, NO RESPONSE, GLUC>160
* DIAB=5: NO RECORD, NO RESPONSE, GLUC 111-121, HBA1C 8.5-9.2
* DIAB=6 IF ONLY HOSP=1 or QUEST=1, WHILE MED=. AND GLUC/HBA1C=NORMAL
* ID=5 and ID=22 have a death record with diabetes
* ID=28 has a medical record with GLUCOSE INTOLERANCE (DIAB=0) ***;
055 INPUT ID META1 P250 INSULIN BOOKMED QMO QYR PMO PYR BOOKMO BOOKYR

```



```

085 DATA VOC73; SET POOLED1; %PERIOD1(4); ***** PERIOD=1973 *****;
086 DATA VOC74; SET POOLED1; %PERIOD1(5); ***** PERIOD=1974 *****;
087 DATA VOC75; SET POOLED1; %PERIOD1(6); ***** PERIOD=1975 *****;
088 DATA VOC76; SET POOLED1; %PERIOD1(7); ***** PERIOD=1976 *****;
089 DATA VOC77; SET POOLED1; %PERIOD2(8); ***** PERIOD=1977 *****;
090 DATA VOC78; SET POOLED1; %PERIOD2(9); ***** PERIOD=1978 *****;
091 DATA VOC79; SET POOLED1; %PERIOD2(10); ***** PERIOD=1979 *****;
092 DATA VOC80; SET POOLED1; %PERIOD2(11); ***** PERIOD=1980 *****;
093 DATA VOC81; SET POOLED1; %PERIOD2(12); ***** PERIOD=1981 *****;
094 DATA VOC82; SET POOLED1; %PERIOD2(13); ***** PERIOD=1982 *****;
095 DATA VOC83; SET POOLED1; %PERIOD2(14); ***** PERIOD=1983 *****;
096 DATA VOC84; SET POOLED1; %PERIOD2(15); ***** PERIOD=1984 *****;
097 DATA POOLED2;
098 MERGE POOLED1 VOC70 VOC71 VOC72 VOC73 VOC74 VOC75 VOC76 VOC77 VOC78
099 VOC79 VOC80 VOC81 VOC82 VOC83 VOC84; BY ID;
099a ***** FOR PERIOD N=N-1 DELETE LINE 082 ('DATA VOC70'), DELETE 'VOC70' *****
099b ***** FROM line 098 (MERGE), AND ADD NEW LINE 101 (VOC1=0 VCL1=0) *****;
100 VOC16=TAEVOC1; VCL16=TAEVOC2; ***** PERIOD=1985 *****;
101 ***** VOC1=0; ***** VCL1=0;
102 ARRAY POC(16) VOC1-VOC16; ARRAY PCL(16) VCL1-VCL16; DO I=1 TO 16;
103 IF POC(I)=. THEN POC(I)=0; IF PCL(I)=. THEN PCL(I)=0;
104 IF CASE=. THEN DO; POC(I)=0; PCL(I)=0; TAEVOC1=0; TAEVOC2=0; END;
105 END;
106 VOCP1=VOC1; VOCP16=VOC16-VOC15;
107 VOCP15=VOC15-VOC14; VOCP14=VOC14-VOC13; VOCP13=VOC13-VOC12;
108 VOCP12=VOC12-VOC11; VOCP11=VOC11-VOC10; VOCP10=VOC10-VOC9;
109 VOCP9=VOC9-VOC8; VOCP8=VOC8-VOC7; VOCP7=VOC7-VOC6; VOCP6=VOC6-VOC5;
110 VOCP5=VOC5-VOC4; VOCP4=VOC4-VOC3; VOCP3=VOC3-VOC2; VOCP2=VOC2-VOC1;
111 VCLP1=VCL1; VCLP16=VCL16-VCL15;
112 VCLP15=VCL15-VCL14; VCLP14=VCL14-VCL13; VCLP13=VCL13-VCL12;
113 VCLP12=VCL12-VCL11; VCLP11=VCL11-VCL10; VCLP10=VCL10-VCL9;
114 VCLP9=VCL9-VCL8; VCLP8=VCL8-VCL7; VCLP7=VCL7-VCL6; VCLP6=VCL6-VCL5;
115 VCLP5=VCL5-VCL4; VCLP4=VCL4-VCL3; VCLP3=VCL3-VCL2; VCLP2=VCL2-VCL1;
*** different from VOC1, VOC2 .. VOC16, VOCP1 ... VOCP16 represent TAEVOC
*** limited to the index period proper, disregarding TAE in former period;
116 DROP C2 LTCA LPCE LTCE LDCE LCIS LDCA LEDC BTCA BPCE BTCE BDCE BCIS
117 BDCA BEDC TCA PCE TCE DCE DCA CIS EDC CHL TTCA TPCE TTCE TDCE
118 TDCA TCIS TEDC I TIN TOUT TSTOP YROUT YRSTOP;
119 DATA NEWTIN; SET RAW; KEEP ID TIN; *** TIN BACK TO ORIGINAL VALUE **;
120 DATA OUT.DIABFILE; MERGE POOLED2 NEWTIN;
*** final diabetes dataset for statistical analysis ***;
*** below follows the list of potential and established cases ***;
*** for clarification of the values in the columns, see lines 55-56 ***;
121 //LISTING DD *
122 1 1 1 1 2 . 83 12 76 11 82 0 111 5.1 1 12 76
123 2 1 1 1 2 2 82 3 85 2 82 0 92 5.1 1 2 82
124 3 1 1 0 . 3 83 3 83 . . 0 181 7.9 1 3 83
125 4 1 0 0 2 2 85 . . 2 85 0 156 6.9 1 2 85
126 5 1 1 1 1 . 82 6 80 . . . . 1 6 80
127 6 1 0 0 . 12 84 . . . . 0 91 6.4 6 12 84
128 7 0 0 0 0 . . . . . 0 126 9.8 3 7 85
129 8 1 1 0 . 1 77 1 79 . . 0 100 5.6 2 1 77
130 9 1 0 0 . 11 75 . . . . 0 95 5.9 6 11 75
131 10 1 1 0 . . 73 4 73 . . 2 226 12.0 1 4 73

```

132	11	0	1	0	.	.	.	10	79	6	10	79
133	12	1	0	0	.	.	78	.	.	.	0	70	6.9	6	6	78	
134	13	1	1	1	1	7	83	4	83	4	84	2	181	10.7	1	4	83
135	14	1	0	.	.	3	78	0	89	7.2	6	3	78
136	15	1	0	0	.	.	81	0	77	6.3	6	6	81
137	16	1	1	1	1	6	72	6	62	.	84	0	213	12.5	1	6	62
138	17	1	1	1	1	12	79	5	78	1	5	78
139	18	1	1	1	2	.	73	3	84	.	73	0	149	6.3	1	6	73
140	19	0	1	0	.	.	.	1	84	.	.	0	98	7.2	6	1	84
141	20	0	1	0	.	.	.	12	77	.	.	0	91	5.8	6	.	.
142	21	0	1	0	.	.	.	12	77	.	.	0	87	5.7	6	.	.
143	22	1	1	1	1	.	70	6	69	1	6	69
144	23	1	1	1	1	.	77	2	79	11	82	3	264	10.7	1	6	77
145	24	1	1	0	2	.	65	.	.	8	65	0	103	7.0	1	8	65
146	25	0	0	0	0	0	112	8.9	5	11	85
147	26	1	0	1	.	.	80	0	198	13.5	1	6	80
148	27	1	0	0	.	.	80	0	95	8.4	6	6	80
149	28	0	0	0	.	.	.	1	82	.	.	0	108	6.9	6	1	82
150	29	1	0	1	1	.	63	.	.	9	63	3	325	9.1	1	9	63
151	30	1	1	0	2	.	78	7	78	.	78	0	197	9.4	1	7	78
152	31	1	1	0	2	8	80	7	84	.	80	0	131	7.3	1	8	80
153	32	1	0	0	.	11	78	0	121	7.5	1	11	78
154	33	1	0	0	.	10	70	0	90	5.9	6	10	70
155	34	1	1	1	2	.	72	5	77	2	85	0	149	7.2	1	6	72
156	35	1	0	0	.	12	78	0	88	5.8	6	12	78
157	36	0	0	0	0	0	169	8.2	4	11	85
158	37	0	0	0	0	0	112	8.4	5	8	85
159	38	1	1	.	.	6	82	8	81	.	.	0	144	7.6	1	8	81
160	39	1	1	1	.	.	.	11	80	1	11	80
161	40	1	1	1	1	2	79	2	78	2	79	0	253	8.6	1	2	78
162	41	0	0	0	0	0	173	7.1	4	9	85
163	42	1	1	0	.	.	.	8	85	.	.	0	161	6.1	1	6	85
164	43	1	1	0	2	.	73	12	73	.	73	0	185	10.3	1	12	73

AUTHORITY: ACT 368 P.A. 1978, AS AMENDED

CEHS-183 (12-88)

"The Michigan Department of Public Health will not discriminate against any individual or group on the basis of race, color, religion, national origin or ancestry, age, sex (or marital status) or handicap."